REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree PhD Year 2005 Name of Author DOSBIM, A.L.

Modelling studies of possible coupling mechanisms between the upper and middle atmosphere

Alison Louise Dobbin

Department of Physics and Astronomy

University College London

This thesis is presented for the degree of
Doctor of Philosophy

June 2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

In recent years, the evidence for a correlation between lower atmospheric behaviour and changes in solar activity has increased. This correlation is thought to be associated with complex coupling mechanisms that link the upper and lower regions of the atmosphere. Dynamical processes are believed to play an important role via the influence of large- and small-scale waves that can lead to the transport of energy, momentum and chemical constituents. Compositional and energetic changes that occur as a direct result of changes in solar energy flux are also thought to contribute to the coupling of atmospheric regions. The production and transport of odd nitrogen species is thought to be particularly important, especially following periods of enhanced geomagnetic activity.

The UCL Coupled Middle Atmosphere - Thermosphere (CMAT) general circulation model was developed in order to provide a tool to investigate two-way coupling between the Earth's lower and upper atmosphere. This model has been updated to include high-resolution empirical solar flux data and current photoabsorption and ionisation cross sections. Improvements have been made to the chemical scheme, and a new spectral gravity wave parameterisation has been implemented

Factors influencing production and loss of odd nitrogen species in the lower thermosphere have been investigated. The representation of nitric oxide in the CMAT model is presented and compared with empirical data. Studies have been carried out to investigate the global production, transport and impact of nitric oxide in the MLT (mesosphere lower thermosphere) region during and following periods of enhanced solar and geomagnetic activity, including a simulation of the 2003 'Halloween super storm'.

The effect of the new gravity wave parameterisation on the zonal and meridional wind structure is also investigated along with its effect on the equatorial local diurnal maximum of atomic oxygen and associated O(^{1}S) 557.7nm green line volume emission rates.

CONTENTS

FIGURE CONTENTS

TABLE CONTENTS

CHAPTER I. BACKGROUND THEORY

1.1 INTRODUCTION

For many years, observed correlations between lower atmospheric properties, the 11-year cycle of solar activity, and variations in geomagnetic activity have been the subject of much debate. The mechanisms through which middle and lower atmospheric parameters may be affected by changes in solar and geomagnetic activity are however still uncertain. For example, measurements of the total solar irradiance at the extremes of the solar cycle show only a very small variation, meaning any relationship between perceived fluctuations in climate and solar energy input is not an obvious one. Detailed modelling and observational studies have revealed a host of dynamical, compositional and energetic processes that are directly impacted by changes in solar radiation and/or auroral energy inputs. These processes may act to enhance and translate changes in solar energy input into the variability observed in climatological properties. An understanding of these natural sources of variability and their influence on the atmosphere is essential if the impact of human activity on the Earth's climate is to be conclusively assessed.

In order to determine the impact of solar variability on the atmosphere, it is necessary to understand the complex coupling mechanisms that connect the upper and lower atmosphere. The work presented within this thesis will explore some of these coupling mechanisms in detail. While studies of climate change are not the focus of this work, it is hoped that improved knowledge of atmospheric coupling processes may assist future work in that area. In this chapter, the basic structure of the atmosphere will be discussed, along with the fundamental processes that govern its behaviour. An introduction of atmospheric modelling is also presented.

1.2 BASIC ATMOSPHERIC STRUCTURE

The vertical structure of the Earth's atmosphere is highly variable in temperature, dynamics and composition. The varying temperature gradient is often used as a way to define different layers or *spheres*. Five main regions exist under this classification, namely the troposphere, stratosphere, mesosphere, thermosphere and exosphere, each

region being characterised by either increasing or decreasing temperature with height (see Figure 1.1). The boundaries between the troposphere, stratosphere and mesosphere are referred to as *pauses*.

In the lowest layer, called the *troposphere*, the temperature decreases with increasing altitude. Shortwave radiation from the Sun is absorbed at the Earth's surface then reemitted upward. As the density of the atmosphere decreases with height, so too does the efficiency with which heat is transported, leading to the decrease in temperature with altitude. The *tropopause* is defined by a temperature minimum and its location changes with latitude and season. At the equator the tropopause is located at roughly 18km altitude where the temperature drops to about 190K. At the poles this temperature minimum reaches 220K at around 8 km altitude. Infrared radiation is the main energy input to the troposphere. This comes primarily from the Earth's surface that absorbs short wavelength solar radiation then re-emits in the infrared. Infrared radiation is also absorbed directly by water vapour. This energy is redistributed by turbulence. The dynamics of the troposphere are driven by pressure gradients and the Coriolis force.

Above the tropopause is the *stratosphere*, a region where the temperature increases with altitude. The *stratopause* is characterised by a temperature maximum of about 270K that occurs at around 50 km altitude. In this region, absorption of solar radiation by ozone (O_3) is the main heating source. The stratospheric ozone also protects the Earth by absorbing harmful solar ultraviolet radiation. A north-south temperature gradient is caused by seasonal differences in hemispheric heating. This gradient, combined with Coriolis forces, leads to strong zonal jets that are eastward in the winter and westward in the summer hemisphere.

The layer above the stratopause is called the *mesosphere* and is another region of decreasing temperature with height. A sharp decrease in ozone concentrations leads to a reduction in ozone heating. Infrared radiative emission, associated with the vibrational relaxation of carbon dioxide (CO_2), provides an important source of cooling in the mesosphere. The upper boundary of the mesosphere, termed the

mesopause, is a temperature minimum that occurs at about 85-100 km. The mesopause temperature can reach as low as 120K in the summer polar regions. Despite large radiative heating rates in the summer, the mesopause is cooler in the summer than in the winter hemisphere. This feature is known as the *mesopause anomaly* and is believed to be the result of dynamical effects that will be described in more detail below.

Above the mesopause is the *thermosphere*, where strong absorption of solar UV and EUV results in a rapid increase in temperature with height. Temperatures are closely related to the level of solar activity and can reach a maximum of 500 to 2000K. Cooling occurs through downward conduction of heat. Above about 150-200 km, molecular heat conduction dominates and the temperature is effectively constant with increasing altitude, up to around 400-600 km. Above this altitude is a region called the *exosphere* in which the density of the atmosphere is so low that neutral particles are unaffected by collisions and move on ballistic trajectories. The fastest particles can actually escape into space.

An alternative way of defining different layers in the atmosphere is by looking at the changes in composition and mixing. Atmospheric density falls off exponentially with height and only 0.1% of the total mass exists above 50km. In the lower atmosphere the major constituents, molecular nitrogen (N_2) and oxygen (O_2), are uniformly mixed by turbulence and make up about 80 and 20% respectively of the total number density. This region is referred to as the *homosphere* and the molecular weight of air varies little with altitude. The top of the homosphere is called the *turbopause* (or *homopause*) and occurs at about 100 km. Above the turbopause, the proportion of atomic oxygen increases, whereas the proportion of molecular species (O_2 and N_2) decrease. In this region called the *heterosphere*, the molecular mean weight of air varies with altitude and constituents are separated by diffusion instead of turbulent mixing.

One final important region of the atmosphere that spans the mesosphere and thermosphere is referred to as the *ionosphere*. Above about 60 km, neutral particles are ionised by solar radiation. At high latitudes, precipitating particles also contribute to

ion production. In the ionosphere, ion and electron densities are sufficiently large as to affect the propagation of radio waves. This region will be described in more detail in section 1.4.5.

1.3 BACKGROUND PHYSICS

1.3.1 HYDROSTATIC EQUILIBRIUM

Under the influence of the Earth's gravitational field, the density of the atmosphere decreases with increasing altitude. Vertical motion is generally small and on large timescales so we can assume the force due to gravity exactly balances the force due to vertical pressure gradient.

If P is pressure at altitude z measured vertically upwards from the surface, g the acceleration due to gravity, and ρ the mass density, the equation of hydrostatic equilibrium can be expressed as follows.

$$\frac{dP}{dz} = -g\rho \qquad 1.1$$

The acceleration due to gravity may be assumed constant since the total depth of the Earth's atmosphere is small compared to its radius.

1.3.2 THE EQUATION OF STATE

The equation of state for an ideal gas of mean molecular mass m and temperature T can be expressed as

$$\rho = \frac{Pm}{kT} \qquad 1.2$$

where k is Boltzmann's constant, ρ the mass density and P the pressure.

Substituting 1.2 into 1.1 and integrating, we get an expression for the pressure P at altitude z

$$P = P_o \exp\left(-\int_{z_o}^{z} \frac{dz}{H}\right) \quad 1.3$$

where P_o is the Pressure at z=0. The scale height H is the increase in altitude required to reduce the pressure by a factor e and is given by

$$H = \frac{kT}{mg} \quad 1.4$$

In the Earth's lower atmosphere, H varies between 6km at T=210K to 8.5 km at T=290K (*Houghton*, [2002]). Above the turbopause where the atmospheric gases are no longer turbulently mixed, the density of each constituent falls off with its own scale height.

1.3.3 THE FIRST LAW OF THERMODYNAMICS

The first law of thermodynamics simply expresses the principal of the conservation of energy. If ΔU is the change in internal energy of a system, Q is the heat added to or absorbed by a system and W is the work done by a system

$$\Delta U = Q - W \quad 1.5$$

1.3.4 THE ADIABATIC LAPSE RATE

If the temperature of a gas is changed adiabatically, it is heated or cooled by compression or expansion without heat being added or taken. When a parcel of air rises in the atmosphere, pressure decreases and the parcel must expand. In expanding it is doing work and therefore uses heat. The result is a lowering of temperature as well as a decrease in pressure and density. When a parcel of air descends in the atmosphere, pressure increases so the gas contracts and undergoes an increase in temperature.

The lapse rate of the atmosphere is defined as the vertical temperature gradient. This is not perfectly adiabatic as differential heating and cooling occur by, for example, absorption of solar radiation.

The dry adiabatic lapse rate Γ_D is defined as the drop in temperature a dry parcel of air would experience if it was lowered or raised adiabatically. Under such conditions, consider the vertical motion of a parcel of air at pressure P, temperature T and of volume V. Assuming hydrostatic equilibrium, the first law of thermodynamics may be expressed as

$$C_V \frac{dT}{dt} + P\frac{dV}{dt} = 0 \qquad 1.6$$

where C_V the heat capacity of air at constant volume and dT/dt is the rate of change of temperature with time.

This equation can be combined with the differential of the equation of state for an ideal gas to give

$$C_V \frac{dT}{dt} = -R\frac{dT}{dt} + V\frac{dP}{dt} \qquad 1.7$$

where R is the gas constant per mole.

By rearranging 1.7 and applying equation 1.1, the following expression for Γ_D can be derived

$$\Gamma_D = -\frac{dT}{dz} = \frac{g}{c_p} \qquad 1.8$$

where z is the vertical coordinate, g is the acceleration due to gravity and c_P is the specific heat capacity at constant pressure. In the Earth's lower atmosphere the specific heat capacity, c_p, has a value of about 1005 J $kg^{-1}K^{-1}$ meaning Γ_D is approximately 10 K km^{-1}.

Consider a column of air in which the lapse rate (or temperature gradient) is less than the dry adiabatic lapse rate. If a parcel of air is displaced upwards from point A to B, it will cool at the dry adiabatic lapse rate. On arrival at point B, the parcel would be colder than the surrounding air and would therefore sink again. This demonstrates a

stable condition. If we consider a column of air in which the lapse rate is greater than the adiabatic lapse rate, on arrival at point B, the displaced parcel will be warmer than the surrounding air and therefore continues to rise. This is an unstable condition. If the lapse rate is perfectly adiabatic and the background atmosphere is in a state of rest, the parcel of air will oscillate vertically due to buoyancy at the *Brunt-Väisälä* frequency.

1.3.5 GEOSTROPHIC BALANCE

In the atmosphere pressure gradients are produced by heating from solar radiation. An element of fluid will move down a pressure gradient and experience a Coriolis force perpendicular to the direction of motion. The Coriolis force acts on any object that is in motion on a rotating frame of reference such that it will move clockwise in the northern hemisphere and anticlockwise in the southern hemisphere.

Newton's second law applied to an element of fluid of density ρ, in the presence of a pressure gradient ∇p can be expressed as

$$\frac{d\mathbf{V}}{dt} = -\frac{1}{\rho}\nabla p + f\mathbf{V} \wedge \mathbf{k} \qquad 1.9$$

where $\mathbf{V}$ is the horizontal wind vector, ∇ is the horizontal del operator. The Coriolis parameter $f = 2\Omega\sin\phi$, where Ω is the Earth's angular rotation rate, ϕ the latitude and $\mathbf{k}$ the unit vector along the axis of rotation.

The first right hand term is the pressure gradient force, the second the Coriolis force. When these two forces balance, such that $d\mathbf{V}/dt=0$, the resulting motion is known as *geostrophic*. The geostrophic velocity is given by

$$\mathbf{V}_g \wedge \mathbf{k} = \frac{1}{f\rho}\nabla p \qquad 1.10$$

The approximation works well at heights where friction is low, and can be used to describe the circulation of wind, parallel to isobars, around centres of high and low pressure. The approximation can also be used to describe the mesospheric jets that arise from a global pressure gradient associated with latitudinal variation in solar

heating. The balance of forces causes the jets to flow westward in the summer hemisphere and eastward in the winter hemisphere as demonstrated in Figure 1.2. In the upper atmosphere (above approximately 100km), ion drag and viscous forces become important and the approximation is no longer valid.

1.3.6 CHEMICAL AND DYNAMICAL TIMESCALES

The abundance of most atmospheric species is determined by a balance between the rates of photochemical production and loss, and the rate of atmospheric transport. The relative importance of dynamical and chemical processes can be examined by evaluating the time constants associated with each process.

The rate of a chemical reaction describes the rate at which the products are formed or reactants disappear. Consider the decomposition of particle A

$$A \rightarrow C + D \quad (k_n) \tag{1.11}$$

The rate is given by

$$\frac{d[A]}{dt} = -\frac{d[C]}{dt} = -\frac{d[D]}{dt} = -k_n[A] \tag{1.12}$$

where k_n is the reaction rate constant (s^{-1} for a unimolecular process) and [N] denotes the concentration of species N (molecules m^{-3}).

If we rearrange this equation we get

$$\frac{d[A]}{[A]} = -k_n dt \tag{1.13}$$

Integration of equation 1.13 gives

$$[A] = [A]_o e^{-k_n t} \tag{1.14}$$

where $[A]_o$ is the concentration of A at time t=0. $1/k_n$ is the chemical lifetime of species A, defined as the time required for the concentration of A to decrease to 1/e of its initial value.

If we wish to examine the lifetime and rate of change of a species in the atmosphere, we must consider all its possible production and loss processes. Consider the following reactions

$$A + B \rightarrow Products \quad (k_1) \tag{1.15}$$

$$A + C + M \rightarrow Products \quad (k_2) \tag{1.16}$$

$$D + E \rightarrow A + Products \quad (k_4) \tag{1.17}$$

Note that the reaction rate constant k_n is in units of m^3s^{-1} for a two body reaction and m^6s^{-1} for a three body reaction. The rate of change of [A] is given by

$$\frac{d[A]}{dt} = -k_1[A][B] - k_2[A][C][M] + k_4[D][E] \tag{1.18}$$

$$= -L[A] + P \tag{1.19}$$

where L represents the sum of the loss processes and P, the sum of the production processes. As for the one body reaction, we can integrate from time t=0 to t and get an expression for [A] at time t

$$[A] = \frac{P - (-[A]_0 L + P)e^{-Lt}}{L} \tag{1.20}$$

The chemical lifetime can be found by substituting $A=A_0/e$ into the above equation

$$\tau = \frac{1}{L} \tag{1.21}$$

If the densities of B, C, D and E are not changing over the time scale considered, and the lifetime is short compared to transport timescales, A is said to be in *photochemical equilibrium*. In this situation, $d[A]/dt=0$ so production is equal to loss and the concentration of A is given by

$$[A]=\frac{P}{L} \qquad 1.22$$

Time constants for dynamical effects on atmospheric constituents can be easily determined if we assume that the number density, n, of each constituent, i, decreases by a factor of e over its scale height H_i (assuming H_i is not variable with height).

$$n_i = n_{i,0}\exp\left(\frac{-z}{H_i}\right) \qquad 1.23$$

where z is the altitude and $n_{i,0}$ is the density of i at a reference altitude.

The vertical gradient of the species can therefore be expressed as

$$\frac{dn_i}{dz} = -\frac{1}{H_i}n_{i,0}\exp\left(\frac{-z}{H_i}\right) = -\frac{1}{H_i}n_i \qquad 1.24$$

Consider the case where changes in the density of a species are due only to vertical winds, w (positive downwards)

$$\frac{dn_i}{dt} = -\frac{d(n_i w)}{dz} = -\frac{w}{H_i}n_i \qquad 1.25$$

Integrating from time t=0 to t we get an expression for the density of species i at time t

$$n_{i,t} = n_{i,0}\exp\left(-\frac{w}{H_i}t\right) \qquad 1.26$$

where $n_{i,0}$ is the density of i at time t=0.

Thus, the density of a species will be $1/e$ of its original value in time H_i/w. This defines the time constant for transport by vertical winds.

Time constants for transport by meridional and zonal winds can be derived in a similar way by replacing H_i by a scale length. When assessing horizontal time constants, care should be taken as the gradients of chemical species are less well known in the horizontal directions. Incorrect assumptions about the gradient of a species can lead to errors in the calculation of dynamical time constants and thus the implied importance of dynamical processes.

For chemically active species the vertical time constant is of the order of days in the upper mesosphere and months in the stratosphere. Similar timescales are estimated for transport by meridional winds. The time constant for transport by zonal winds is of the order of days throughout the middle atmosphere (*Brasseur and Solomon* [1986]).

One final source of vertical transport to be considered here is referred to as *eddy* or *turbulent diffusion* and describes the transport of mass by small scale eddies (where *eddies* refers to fluctuations about the zonal mean). The time constant for transport by vertical eddy diffusion is given by H^2/K_z (*Brasseur and Solomon* [1986]), where K_z is the vertical eddy diffusion coefficient.

As mentioned at the start of the section, comparison of chemical and transport timescales can help us understand which processes determine constituent concentration at a particular point in the atmosphere. To illustrate this point, consider the case where the transport timescale of a species is much larger than the chemical timescale. The concentration of that species will be predominantly determined by chemistry and transport may be neglected to a first order approximation. Alternatively, if the transport timescale is much smaller than the chemical timescale, the constituent will be well mixed. When the two timescales are comparable, the distribution of the species will be dependent on both dynamics and chemistry.

Note that when a species is in photochemical equilibrium, the concentration of that species may be affected through temperature changes associated with dynamical

effects, or through the transport of longer-lived constituents that contribute to its chemistry. Thus transport processes can indirectly affect the species concentration and cannot be neglected.

1.3.7 THE ABSORPTION OF SOLAR RADIATION

At any point in the atmosphere, the intensity of solar radiation can be calculated from the Lambert-Beer exponential absorption law.

$$I(\lambda) = I_{\infty}(\lambda)\exp[-\tau(\lambda)] \qquad 1.27$$

where $I(\lambda)$ is the intensity of radiation at wavelength λ at any point in the atmosphere and $I_{\infty}(\lambda)$ is the irradiance at 1AU. The optical depth, $\tau(\lambda)$, specifies the attenuation of solar irradiance by the atmosphere. At an altitude of z_0, the wavelength dependent optical depth is given by

$$\tau_{z_0}(\lambda) = \sum_j \sigma_j(\lambda) \int_{z_0}^{\infty} n_j(z) \sec\chi dz \qquad 1.28$$

where $\sigma_j(\lambda)$ represents the wavelength-dependent absorption cross sections of species j and $n_j(z)$ is the height profile of the species concentration. $sec\chi dz$ represents the infinitesimally small volume through which the radiation travels, at an angle χ to the vertical, as illustrated in Figure 1.3

The angle at which solar radiation penetrates the atmosphere is dependent on local time, season and latitude such that the cosine of the solar zenith angle is given by

$$\cos\chi = \cos(\phi)\cos(\delta)\cos(H) + \sin(\phi)\sin(\delta) \qquad 1.29$$

where ϕ is the latitude, δ the seasonally dependent solar declination angle, H is the hour angle which is 0° for local noon. For solar zenith angles >75° (sunrise or sunset), the curvature of the Earth makes the atmospheric column content a complicated function of solar zenith angle. In such cases, the Chapman grazing incidence function, $Ch(\chi)$, replaces the solar zenith angle. This function represents the ratio of the total

amount of absorbant along the oblique angle (χ with respect to the vertical) versus the total amount of absorbant in the vertical.

1.3.8 RATES OF PHOTOIONISATION AND DISSOCIATION

Photoionisation is the principal mechanism through which the ionosphere is created. The photoionisation rate (in ions $cm^{-3}s^{-1}$) of a species j at altitude z is given by

$$J_z^i(j,l) = n_z(j)\int_{\lambda th}^{0} I_z(\lambda)b_j^i(\lambda,l)\sigma_j^i(\lambda)d\lambda \qquad 1.30$$

where $\sigma_j^i(\lambda)$ is the ionisation cross section. The branching ratios $b_j^i(\lambda,l)$ specify the production of ions in electronically excited or ground states. For each constituent,

$$\sum_l b_j^i(\lambda,l) = 1 \qquad 1.31$$

Equation 1.30 must be integrated from the ionisation threshold wavelength, λ_{th}, for the production of ions in their ground state, to the shortest wavelength at which photon flux contributes to ion production. In modelling it is common practice to specify total photoionisation cross sections and branching ratios for specific wavelength intervals.

Photodissociation of O_2 accounts for O being the major species in the thermosphere above 200km. The photodissociation rate (in $cm^{-3}s^{-1}$) of a species j at altitude z is

$$J_z^d(j) = n_z(j)\int_{\lambda_d}^{\lambda_i} I_z(\lambda)\sigma_j^d(\lambda)d\lambda \qquad 1.32$$

where $\sigma_j^d(\lambda)$is the photodissociation cross section. The integration is carried out from the wavelength below which photons play a part in the process, λ_d, to the ionisation limit, λ_i, where photon absorption leads almost entirely to ionisation.

1.3.9 LOCAL THERMODYNAMIC EQUILIBRIUM

In the Earth's lower atmosphere it is useful to assume a state of *Local Thermodynamic Equilibrium* (LTE). In this region, excitation and de-excitation by collisions are

sufficiently rapid that they dominate over radiative processes such as photon absorption, spontaneous emission and collisionally activated emission.

When there is significant net energy gain or loss from radiative processes or when rates of excitation or de-excitation by collisions are comparable with rates of excitation or de-excitation by absorption or emission, the condition no longer applies and the atmosphere is in a state of non local thermodynamic equilibrium (non-LTE). For a full discussion of LTE and non-LTE processes see *Lopez-Puertas and Taylor* [2001].

1.4 THE MIDDLE-UPPER ATMOSPHERE SYSTEM

1.4.1 HEATING

Studies of the energy balance of the middle and upper atmosphere must include the consideration of several key physical and chemical processes. Absorption of solar radiation, the storage and release of chemical energy, energy transport, dissipation of wave disturbances, auroral heating and cooling by radiative emissions all contribute to the energy budget of the atmosphere.

Knowledge of the diverse energy inputs from the sun is obviously essential in understanding the energy budget of the atmosphere. High energy, short wavelength radiation is mostly absorbed above 100km, while long wavelength visible and infrared radiation can penetrate to the surface. At wavelengths over 100nm, the solar spectrum is subdivided into regions of absorption by primary absorbing species, O_2 and O_3. The depth of penetration of solar radiation and the key absorbing species are illustrated in Figure 1.4. The key wavelength ranges of solar radiation that are absorbed in the atmosphere are outlined in Table 1.1.

In the stratosphere absorption of solar ultraviolet radiation by ozone in the Hartley (203-305nm) and Huggins (305-397nm) bands provides the principal source of heat. The heating rate can reach a maximum of about $12Kday^{-1}$ near the stratopause but is dependent on the amount of ozone present.

The principal absorbers in the mesosphere are O_2, O_3 and CO_2, which absorb wavelengths ranging from Lyman alpha (121.5nm) in the UV, to the mid infrared

(4.3μm). As the concentration of ozone drops off with increasing height, so does the heating associated with the absorption of radiation in the Hartley and Huggins bands. Most band energy absorbed by O_3 and O_2 in the mesosphere does not initially cause heating but is redistributed as chemical potential energy or internal energy of electronically or vibrationally excited species that are produced by photolysis reactions. Internal energy can either be quenched to heat through collisions, or it may be radiated in the form of airglow, reducing the amount of energy available for heat. The heating rate therefore depends strongly on how and where this chemical and internal energy is deposited.

One significant transport of absorbed solar energy occurs through the dissociation and eventual recombination of O_2. Atomic oxygen, created by the dissociation of O_2, stores much of the energy absorbed in the photolysis reaction as chemical energy. Above 75km atomic oxygen has a lifetime of over one day and thus is transported away from the original reaction site. When recombination occurs, the stored chemical energy is released as thermal energy.

In parts of the middle atmosphere, specifically between about 70 and 95km, heating due to exothermic chemical reactions between neutral constituents is comparable to and can exceed that due to direct solar heating. In the mesosphere, a local temperature maximum arises from the recombination of O and O_2 to form ozone and quenching of $O(^1D)$ created through dissociation of ozone. Reactions involving odd hydrogen are key, particularly that between atomic hydrogen and ozone which may be the single largest heat source between 83 and 95km altitude. For a detailed study of the role of exothermic reactions in the energy budget of the mesosphere see *Mlynczak and Solomon* [1993].

Above 80km, the majority of solar photons at wavelengths below 200nm are absorbed due to the large absorption cross-section of atmospheric gases. Direct solar heating by absorption in the Schumann-Runge bands (175 to 200nm) by O_2 begins to play a dominant role in the lower thermosphere. Radiative heating continues to increase with increasing altitude as O_2 absorbs solar radiation in the Schumann-Runge continuum

(130-175nm). A maximum heating rate of 30-40Kday^{-1} is reached at about 150km. The spectral distribution of the absorption cross section of O_2 is shown in Figure 1.5.

Heating due to EUV (10 to 100nm) dominates in the 150-300km altitude region. Maximum EUV heating rates at solar minimum and maximum can vary between about 500 and 2000 Kday^{-1}, leading to dramatic changes in temperature changes over the solar cycle. The absorbing species, O, O_2 and N_2 regularly undergo photoionisation by solar photons in this altitude region. If the incident photon is sufficiently energetic, the photoionisation event can result in the production of energetic photoelectrons. These fast photoelectrons are ejected, carrying a portion of the initial photon energy with them. The rest is transformed to chemical energy of the ion product. Local heating occurs when the fast photoelectrons collide with electrons in the background ionosphere. Subsequent collisions between these electrons and neutral particles cause thermal energy to be transferred to the neutral gas. Direct collisions between energetic photoelectrons and neutrals can also result in energy transfer. If the photoelectrons are sufficiently energetic, collisions with neutrals may lead to either further ionisation, or excitation of internal atomic levels of an atom or molecule in the gas. Collisions between these excited particles and neutrals can result in deactivation, the excess energy appearing as local thermal energy of the particles. Alternatively the energy can be radiated as airglow emissions.

The chemical energy of the ion can be regained by recombination with an electron, or transferred to other ions through a series of charge transfer reactions. Exothermic chemistry can play a role in redistributing the energy locally. Some energy is transported, specifically by atomic oxygen which is created through chemical reactions following the ionisation event. The O atoms travel to lower altitudes where the increased particle density enables a three-body recombination reaction to proceed. This reactions takes the form $O + O + M = M + O_2$, where M is a third molecule of gas. In this way, stored thermal energy is transported away from the location of the photoionisation event then released below 90km.

Large amounts of energy from the solar wind are deposited in high latitude auroral

regions through Joule heating and particle precipitation. Joule heating is caused by collisions between neutrals and energetic ions that have been accelerated by the high latitude electric field. Particle precipitation causes heating by collisions between the neutral gas and high-energy ions and electrons that have been accelerated along the Earth's magnetic field lines into the polar thermosphere. The strength of the magnetic field controls the depth of penetration of precipitating electrons. Magnetospheric perturbations cause large variations in the flux of precipitating electrons and thus, in the subsequent thermospheric heating.

Finally, the role of dynamical processes in the energy budget of the atmosphere should be considered. For example, the formation of mesospheric inversion layers, temperature enhancements sandwiched by cooling regions, is thought to be related to gravity wave-tidal interaction (*Liu and Hagan* [1998]). *Liu et al.* [2000] showed that the dissipation of gravity waves in the mesosphere could lead to local heating rates as large as $10 K day^{-1}$. *Medvedev and Klaassen* [2003] derived terms describing the thermal effects of gravity waves on the mean flow and concluded that the combined effect of saturating gravity waves is to produce both differential heating and cooling. Unlike the heating or cooling due to chemical or radiative processes, dynamical sources and sinks are very difficult to measure by conventional means. However, with accurate information about chemical and radiative sources and sinks, measurements and modelling studies can be used to infer the net effect of dynamical processes.

Feature	**Wavelength Range**	**Principal Atmospheric Absorber**
Middle infrared bands	2.0, 2.7 and 4.3μm	Absorption by CO_2 in the mesosphere
Atmospheric band	762nm	Absorption by CO_2 in the mesosphere
Chappuis Band	397-850nm	Absorption by O_3 in the troposphere
Huggins Band	305-397nm	Absorption by O_3 in the stratosphere and troposphere
Hartley Band	203-305nm	Absorption by O_3 in the stratosphere and mesosphere
Hertzberg continuum	200-240nm	Absorption by O_3 and O_2 in the mesosphere
Schumann-Runge Bands	175-200nm	Absorption by O_2 in the lower and middle and thermosphere
Schumann-Runge Continuum	130-175nm	Absorption by O_2 in the middle thermosphere
Lyman alpha	121.5nm	Absorption by O_2 in the mesosphere. Photoionisation of NO lower thermosphere.
FUV	100 - 200nm	Absorption by and photoionisation of O, O_2 and N_2 in middle and upper thermosphere. Photodissociation of O_2 and N_2.
EUV	10 - 100nm	Absorption by and photoionisation of O, O_2 and N_2 in middle and upper thermosphere. Photodissociation of O_2 and N_2.

Table 1.1 Primary absorbers and wavelength ranges of solar radiation absorption from the stratosphere to thermosphere. Adapted from *Beig et al.* [2003].

1.4.2 COOLING

Radiative emission dominates cooling in the stratosphere and mesosphere where the thermal structure arises from an equilibrium between absorption of UV radiation and emission in the infrared. Cooling is accomplished primarily through emission by carbon dioxide at 15μm, although in the lower mesosphere 9.6μm emission by ozone and 80μm emission by water vapour also play a part. A minor contribution also comes from downward turbulent heat conduction. Figure 1.6 shows the vertical distribution of daily heating and cooling rates due to absorption and emission by the major absorbing species.

Below about 70km, collisions between particles are sufficiently frequent that a state of local thermodynamic equilibrium (LTE) can be assumed. At higher altitudes we must take into account non-LTE processes such as radiative absorption between energy levels and energy transitions due to collisions with other particles. Vibrational deactivation of the $CO_2(0110)$ state through collisions with atomic oxygen gives rise to a maximum in CO_2 cooling at the mesopause. Consideration of these complex processes means that calculation of cooling rates in the mesosphere is non trivial.

Above around 140km, cooling by CO_2 emission drops off and downward molecular heat conduction becomes the primary cooling mechanism. Between 120 and 200km, cooling by 5.3μm non-LTE radiation from NO is the dominant radiative cooling mechanism which contributes appreciably to the total cooling rate in the middle thermosphere. 63μm infrared radiation from the fine structure of atomic oxygen is the dominant radiative loss process in the upper thermosphere but it plays a minor role in the overall thermal balance. Figure 1.7 and Figure 1.8 show the global mean, total and component heating and cooling rates in the mesosphere and thermosphere as calculated by *Roble* [1995].

1.4.3 THE MESOPAUSE ANOMALY

If the middle atmosphere were in radiative equilibrium, one would expect the atmospheric temperature to closely follow the distribution of solar heating. Under such conditions, the maximum temperature at a given altitude in the mesosphere should

occur at solstice over the summer pole where the sun shines for 24 hours a day. A temperature minimum would be expected during the polar night where radiatively determined temperatures are considerably lower than those in the summer hemisphere. This is demonstrated in Figure 1.9 (*top*), which shows the zonal mean temperature profile calculated assuming radiative equilibrium (*Geller* [1983]). In reality however, the highest mesospheric temperatures are found over the winter pole, and the lowest over the summer pole. Observed summer mesopause temperatures are in fact lower than those found anywhere else in the atmosphere. This phenomenon is often referred to as the *mesopause anomaly*. The temperatures in the summer mesopause are so low that they can cause *noctilucent*, or *polar* mesospheric clouds, cloud formations thought to be composed of small ice-coated particles.

The origin of the observed mesospheric temperature structure can be more easily understood if we consider the difference between calculated and observed zonal mean wind profiles. Figure 1.9 (*bottom*) shows the calculated zonal mean wind profile associated with the radiative equilibrium temperature profile. The calculated wind profile assumes geostrophic balance and therefore predicts a wind profile that is dominated by zonal mesospheric jets, as described in section 1.3.5. Figure 1.10 shows a mean zonal wind climatology measured by the upper atmosphere research satellite (UARS) (*McLandress et al.* [1996a]). The UARS data shows a *closure* of the zonal jets at about 80km with a reversal in direction at higher altitudes. These plots demonstrate that the middle atmosphere temperature and wind structure cannot be described by geostrophic balance alone.

The presence of an inter-hemispheric circulation at solstice was first suggested by *Murgatroyd and Singleton* [1961] and *Haurwitz* [1961]. As air rises in the summer mesosphere it is adiabatically cooled. In the winter hemisphere, downwelling air is adiabatically heated as it descends. A summer to winter circulation is thus set up in the region of the mesospause. In order to balance the Coriolis force associated with such a meridional circulation, a westward drag is needed in the mesosphere of the winter hemisphere, and an eastward drag in the mesosphere of the summer hemisphere. Inclusion of a meridional drag term, F, in the zonal horizontal momentum equation

results in a calculated zonal wind profile that reproduces this circulation and the closure of the zonal jets.

$$\frac{d\mathbf{V}}{dt} = -\frac{1}{\rho}\nabla p + f\mathbf{V} \wedge \mathbf{k} - F \qquad 1.33$$

Houghton [1978] was the first to suggest that dynamical forcing by small-scale oscillations called gravity waves might provide the required drag. Modelling work by *Lindzen* [1981] and *Matsuno* [1982] showed that if vertically propagating waves were selectively absorbed as they travelled from the troposphere to the mesosphere, more westward waves should be present in the winter mesosphere and more eastward waves in the summer mesosphere. It is now accepted that dissipation or *breaking* of these gravity waves provides the dynamical forcing required to drive the observed circulation.

1.4.4 STRATOSPHERIC CIRCULATION AND PLANETARY WAVES

As in the mesosphere, the thermal structure of the stratosphere is also strongly influenced by dynamical processes. During solstice, the summer stratosphere is not very far removed from radiative equilibrium. The winter hemisphere however, is much warmer than it would be if it were controlled by radiative processes alone. The warm winter hemisphere is a result of the *Brewer-Dobson* circulation (*Brewer* [1949], *Dobson* [1956]), as illustrated in Figure 1.11. Air rises in the tropics, moves poleward, then returns to the troposphere at mid-latitudes. The circulation is caused by planetary wave dissipation in the extra tropical winter stratosphere.

Planetary waves are stationary, or slowly westward propagating oscillations that are generated in the troposphere and propagate up into the middle atmosphere. They arise from background wind interactions with topographic and thermal features, such as large landmasses and land-ocean heating contrasts. *Charney and Drazin* [1961] showed that these waves are not able to propagate under conditions of westward flow as they cannot move through regions where the background wind is equal to the planetary wave horizontal phase speed. Thus significant planetary wave activity in the stratosphere is limited to areas where the zonal flow is eastward, such as in the winter

hemisphere. Waves propagate upwards in the winter hemisphere then dissipate in the middle and upper stratosphere, as illustrated in Figure 1.12.

Planetary wave dissipation can occur through two processes, namely radiative damping or wave breaking. Radiative damping refers to the process of thermal dissipation in which radiative processes lessen the temperature gradients associated with the wave. The timescales for this process are sufficiently long that it is unlikely to play a major role in controlling the amplitude of the planetary waves. Wave breaking occurs in regions where the planetary wave velocity is the same as the velocity of the mean flow. This is thought to be the dominant process by which planetary waves dissipate in the stratosphere, *Shepherd* [2000]. For stationary planetary waves, which dominate the observed spectrum, wave breaking occurs in the subtropics of the winter hemisphere where the zonal wind speed is zero.

When planetary waves break in the region of the wintertime stratospheric polar night jet stream (a strong eastward jet that exists along the polar night terminator), they deposit their westward momentum thus decelerating or displacing the jet stream. This allows warm mid or even low latitude air to enter the cold polar region, resulting in *stratospheric sudden warmings*. In order to regain thermodynamic equilibrium, radiative cooling occurs in the polar winter stratosphere. The cooled air sinks and by mass conservation, is replaced by warm mid and low latitude air that rises from lower altitudes. Thus the equator to pole Brewer-Dobson circulation is set up, allowing the transport of heat, material and momentum from low to high latitudes. Without this circulation, the winter stratospheric pole would be considerably colder than it actually is.

Relative to the southern hemisphere, the northern hemisphere winter experiences greater momentum deposition by planetary waves. This results in a stronger poleward branch of the Brewer-Dobson circulation which in turn, leads to stronger downwelling over the pole and thus warmer polar temperatures. This hemispherical asymmetry in planetary wave activity arises because planetary waves are primarily forced by topography and land-sea contrasts which are much stronger in the northern hemisphere

than in the southern hemisphere. The northern winter hemisphere therefore experiences stronger planetary wave forcing than that of the southern winter.

1.4.5 THE IONOSPHERE

The existence of an electrically conducting atmospheric layer was first suggested in the 19th Century in response to observed diurnal variations in the terrestrial magnetic field. In 1901, *G. Marconi* transmitted the first radio signals from Europe to America. *Kennelly* and *Heaviside* later suggested that the transmitted radio waves were being reflected from a conducting layer near 80km altitude. This conducting layer has now been named the *ionosphere*. Ongoing investigations of the region have revealed that the ionosphere plays a major role in atmospheric dynamics, energetics and composition.

The upper and lower limits of the ionosphere are not well defined, the lower limit occurring around 50 to 60km. Above about 1000km, the ionosphere merges into the magnetosphere, an area around the Earth where ion motion is controlled entirely by the Earth's magnetic field. The layers of the ionosphere are defined by the mean vertical profile of electron density as illustrated in Figure 1.13.

The D region is located between around 60 and 85km. Ionisation in this layer is dominated by photoionisation of nitric oxide (NO) by solar Lyman alpha radiation (121.6nm). Solar Lyman alpha radiation is able to penetrate deep into the atmosphere because it is situated in an *optical window*, a spectral region where absorption by the upper atmosphere is quite weak. High-energy galactic cosmic rays also contribute to ion creation in the lower part of the D region. Ionisation by cosmic rays decreases with increasing solar activity as the intense solar wind pushes the cosmic rays away from the solar system.

Above the D region is the E region, created primarily through photoionisation of molecular nitrogen and oxygen by solar ultraviolet radiation, and atomic oxygen by solar X-rays and Lyman beta radiation (102.6nm). N_2^+ is very short lived because of a rapid charge exchange reaction with oxygen. O_2^+ is long lived and is present in large quantities, along with NO^+ that is created by a reaction between N_2^+ and atomic

oxygen. NO^+ and O_2^+ are therefore the dominant ions in this region which extends to about 130km.

The F region begins above the E region and is divided into the F_1 and F_2 layers. O^+ is the dominant ion, created through ionisation of atomic oxygen by solar EUV radiation (10-100nm). Molecular nitrogen is also ionised by the incident EUV radiation. As the flux of solar EUV is highly variable, so too is the location and size of the electron density peak which reaches a maximum in the F region (Figure 1.14). A typical peak electron density value for midday conditions would be $10^{12}m^{-3}$, occurring between 250 and 300km.

In the D, E and F_1 layers of the ionosphere, the lifetime of ions is short compared to the transport timescale, such that the concentration of charged particles is controlled by a photochemical equilibrium between production and loss. Since the production of charged particles is directly related to the flux of ionising solar radiation, the night-time ionosphere looks very different to that of the day. In the D region, electrons disappear almost entirely on the night side of the globe as charged particles recombine. The electron densities of the E and F_1 regions are also greatly reduced. Above the F_1 region, recombination timescales are increased and processes such as transport by molecular diffusion begin to play a role. The F_2 region is present on the day and night sides of the globe. Above the F_2 peak, electron density decreases rapidly. At altitudes over about 600km, ions of helium then hydrogen dominate. At these altitudes the ions are aligned with the Earth's magnetic field and thus, no longer form horizontal layers.

An appreciation of the photoionisation and absorption cross-sections of atmospheric constituents can assist us in understanding how solar energy is translated into atmospheric thermal and chemical energy.

An important feature of ionisation by high-energy photons, such as hard X-rays, is that once a neutral particle has absorbed the photon, the photoelectrons that are produced have sufficient energy to ionise other particles. We must therefore distinguish between primary and secondary photoelectrons. The number of secondary ion pairs produced is

related to the initial photon energy and can be large enough to significantly affect ionospheric structure.

1.4.6 NEUTRAL-ION INTERACTIONS

The ionosphere plays a major role in the dynamics and energetics of the thermosphere. As previously mentioned, the ions and neutrals collide with each other and thus exchange energy and momentum. The relative influence of neutrals upon charged particles and vice versa depends upon the ratio of *collision frequency* to *gyro frequency*. The *ion-neutral* collision frequency is the rate at which ions encounter neutral particles and can be expressed as $\nu_{in}=nK_{in}$, where n is neutral concentration. K_{in} is the ion-neutral collision rate coefficient which describes the transfer of momentum between ions and neutrals. Note that the ion-neutral collision frequency and the neutral-ion collision frequency are not the same. The latter describes the rate at which neutrals encounter ions and is smaller than the former since ion densities are typically much smaller than neutral densities in the thermosphere.

The gyro frequency relates to the fact that in the presence of a magnetic field, charged particles undergo gyroscopic motion around the field lines. The gyro frequency of a charged particle i is given by $\omega_i=Bq/m$ where B is the magnetic field strength, q is the particle charge, and m the particles mass.

If the gyrofrequency ω_i is much larger than the ion-neutral collision frequency ν_{in}, charged particles will be largely unaffected by the neutral wind and will continue to travel around the magnetic field lines. If the charged particles have a sufficiently large velocity relative to that of the neutral particles, and are present in sufficient quantities, the ions can actually drive the neutral wind. If the ion-neutral collision frequency is much larger than the gyrofrequency, the motion of the ions will be controlled by the neutral wind.

In the lower ionosphere (around 70km) the collision frequencies of ions and electrons are about $10^5 s^{-1}$ and $10^7 s^{-1}$ respectively (*Ratcliffe* [1972]). Ion motion is controlled by the neutral wind. As altitude increases, the ion and electron collision frequencies decrease such that at 300km typical values are $0.5 s^{-1}$ and $10^3 s^{-1}$. The gyro frequencies

of ions and electrons however remain approximately constant over this height range (about $160s^{-1}$ for ions and $8x10^{6}s^{-1}$ for electrons). At a point where a balance is reached between gyro and collision frequencies, ion motion is influenced by both the magnetic field and the neutral wind. In the F region high ion velocities result in a neutral wind that is strongly influenced by ion motion. Note that the velocity of ions is highly variable over latitude and altitude.

The presence of free electrons and ions in the ionosphere means it has a high electrical conductivity. Particle motions are strongly influenced by the geomagnetic field meaning the conductivity is anisotropic in height and direction. Electric currents can arise at low and mid latitudes through the action of the neutral wind that exerts a force on neutral and charged particles via collisions. This occurs in the E region at a point where the ion gyrofrequency is smaller than the collision frequency but the electron gyrofrequency is larger than the collision frequency. Under these conditions, ions move with the wind but electrons move across it under the influence of the geomagnetic field. This movement of electrons and ions constitutes an electric current. Horizontal electric fields are created by the separation of particles of opposite charge, which in turn affects the motion of ions and electrons.

The electric field pattern of the E region is reproduced in the F region where electron and ion gyrofrequencies are much greater than the collision frequency. This leads to plasma drifts where electrons and ions move together.

An electric field is also mapped down into the high latitude F-region from the magnetosphere above. This field is generated as the solar wind causes circulation of the magnetospheric plasma. Plasma movement far out in the magnetosphere is transferred along the moving magnetic field lines to the high latitude ionosphere. At high altitudes the collision frequency is less than the gyro frequency and charged particles follow the lines of moving force. At altitudes of about 140km, collisions become more frequent such that ion motion is no longer controlled by the magnetic field. Electrons continue to be controlled by the magnetic field until a height of around 80km. There is thus a region of charge separation where polar currents flow. Current

flows over the poles towards the sun and returns in the auroral zone. The electric field is directed across the magnetotail from the dawn to dusk side. Under average conditions, the cross cap potential is about 50kV.

At high latitudes, charged particles accelerated in the electric field, collide with neutral particles. These accelerated collisions result in Joule heating which plays a major role in the energy budget and dynamics of the thermosphere.

1.4.7 GEOMAGNETIC ACTIVITY AND PARTICLE PRECIPITATION

The interplanetary magnetic field (IMF) is an extension of the magnetic field of the sun that is 'frozen in' to the solar wind. The concentration and energy spectrum of particles entering the Earth's atmosphere is strongly affected by the orientation of the IMF relative to the geomagnetic field. When the IMF has a strong southward component, magnetic reconnection can occur, leading to a coupling between the magnetosphere and the solar wind. Solar wind plasma can thus enter the magnetosphere system and be propagated along the Earth's magnetic field lines into the atmosphere. Figure 1.15 shows the basic structure of the magnetosphere.

Figure 1.16 illustrates the dominant behaviour of the Earth's magnetic field under conditions of northward and southward IMF. Magnetic field lines are '*closed*' when the IMF is northward, and '*open*' under conditions of southward IMF, allowing the two fields to merge. Note that the details of magnetic reconnection processes are complex and only an outline of the subject has been provided here. The finer details of this topic are the subject of much ongoing research, and are beyond the scope of this discussion.

Energetic charged particles that enter the atmosphere at high latitudes provide an important energy input to the middle and upper atmosphere by heating and ionising neutral particles. In the polar regions where the Earth's magnetic field lines are open, particles that enter the atmosphere are generally of lower energy than those that enter the atmosphere in the auroral zones. Here, particles with typical energy 1-10keV precipitate into the atmosphere from the magnetospheric plasmasheet and cause spectacular optical displays know as the aurora borealis and aurora australis. Particles

with energies of a few MeV are observed in the equatorward cusp of the auroral oval and are able to penetrate deep into the atmosphere.

As with high energy X-rays, high-energy particles produce substantial amounts of secondary electrons with energies from 10 to 100eV. These are responsible for a considerable fraction of the energy transfer from the primary particle to the atmosphere.

During large solar flares or coronal mass ejections (CMEs), large amounts of high-speed plasma, containing protons of 10-300MeV, are injected into the solar wind. On arrival at the magnetopause, if the solar wind has a southward magnetic field, the energy and particles will be effectively coupled into the Earth's atmosphere and a geomagnetic storm will commence. When the storm occurs, particle precipitation increases and the auroral oval expands to lower latitudes than normal. These events can produce intense ionisation in the Earth's D region, particularly at high latitudes, and change the composition of the neutral atmosphere.

A measure of geomagnetic activity is the planetary 3-hour K_p index (*Bartels* [1949]). The index is obtained from the mean value of the disturbance levels in the north-south geomagnetic field, observed at 13 stations. Levels at each station are determined by measuring the range of disturbance, relative to an assumed quiet-day curve, during three-hourly time intervals. The range is then converted into a local K index taking the values 0 to 9 according to a quasi-logarithmic scale. The three-hourly a_p index is a linear scale, directly related to the K_p index. A_p is a daily index obtained from the average of the a_p indices for that day.

1.4.8 GRAVITY WAVES

Gravity waves are small-scale perturbations in wind, temperature, and density where the oscillation is transverse to the direction of propagation. They are generated in the troposphere by a variety of mechanisms including the passage of wind over surface topography, and meteorological phenomena such as storm convection and wind shear. As the waves propagate up into the middle atmosphere, their amplitudes will increase exponentially. This can be understood by considering the conservation of kinetic

energy density $(1/2)\ \rho\ |V|^2$, where ρ is the density of gas at height z in the atmosphere and $|V(z)|$ the amplitude of the vertical velocity component. If we define c as a constant, we can say that $|V(z)| = c\rho^{-1/2}$. Since density varies as $\rho_0 exp(-z/H)$ where ρ_0 is density at height z=0 and H is scale height, it follows that $|V(z)| = (c/\rho_0)\ exp(z/2H)$, i.e. wave amplitude grows exponentially with height.

Buoyancy forces act to restore the oscillation until some altitude is reached where the local temperature perturbation induced by the wave produces a superadiabatic lapse rate. At this point the atmosphere is unstable and the wave begins to dissipate or break. A cascade of turbulence is triggered and energy and momentum are deposited into the mean wind flow.

Lindzen [1981] and *Weinstock* [1982] assumed that above the breaking level, turbulent diffusion arising from wave breaking prevents the wave amplitude from growing any larger. Using this assumption, it is possible to derive expressions for turbulent diffusion and momentum deposition (wave drag), both of which are strong functions of $\overline{u} - c$, the difference between the zonally averaged wind, $\overline{u}$, and the horizontal phase speed of the wave, c. At an altitude where the horizontal phase speed of the wave is equal to the background wind speed, i.e. $\overline{u} - c = 0$, the wave can no longer propagate and is completely absorbed. Strong momentum deposition occurs between the breaking level and this upper limit.

As discussed in section 1.4.3 the zonal mean forces arising from gravity wave breaking cause the reversal of the mid latitude zonal mean jets. Gravity waves in the mesosphere drive the meridional circulation that causes the observed latitudinal temperature gradient to be the reverse of that expected in the absence of wave driving (i.e. the summer mesopause is colder than the winter). Gravity wave dissipation also causes mixing effects on trace constituents. *Hickey* [1994], suggested that the time averaged effect of gravity waves on chemical exothermic heating can lead to heating rates that are greater than those suggested by *Mlynczak and Solomon* [1991, 1993]. Studies suggest that gravity waves can lead to changes in the thermal structure of the

middle atmosphere through wave dissipation and through wave-wave interactions (*Liu and Hagan* [1998], *Liu et al.* [2000], *Medvedev and Klaassen* [2003]).

Gravity wave representation in numerical models will be discussed further in Chapter 2 but it is worth noting here that the scale of general circulation model coordinate grids is typically too large to explicitly resolve these small-scale oscillations. Applying finer grids in order to resolve these waves incurs prohibitive computing overheads. Instead, several physically justifiable parameterisations have been developed which require specification of a gravity wave source spectrum in the lower atmosphere as input. *McLandress* [1998] has reviewed four of the most commonly used gravity wave parameterisations. For a review of gravity wave dynamics and effects in the middle atmosphere see *Fritts and Alexander* [2003].

1.4.9 TIDES

Atmospheric solar tides are planetary-scale oscillations in winds, temperature and pressure with periods that are harmonics of a 24-hour day. These tides play an important role in the dynamics, energetics and composition of the middle and upper atmosphere. Tides generated by periodic heating of the atmosphere are termed *thermal tides* and migrate westward with the apparent motion of the sun. Tides arising from gravitational forcing through the moon and sun are described as *Gravitational tides.* Thermal tides dominate over gravitational tides in the Earth's atmosphere.

Thermally driven tides are primarily generated by heating through absorption of solar IR radiation by water in the troposphere, absorption of UV by ozone in the stratosphere and by EUV absorption by molecular oxygen and nitrogen in the thermosphere. Tides generated in the lower and middle atmosphere have components which are able to propagate vertically into the thermosphere. Observations of tidal amplitudes in the mid to high latitude thermosphere reveal a dependency on geomagnetic activity. This is because oscillations can also be generated by ion-neutral interactions such as ion drag and Joule heating. Modelling studies by *Müller-Wodarg et al.* [2001] suggest that these oscillations are comparable in magnitude and, at high latitudes, stronger than the upward propagating tides.

Thermally driven tides arise as the variation in day-night heating generates Fourier components which are subharmonics of a solar day. The most prominent tides found in the terrestrial atmosphere are the diurnal (24h), semidiurnal (12h) and terdiurnal (8h). Tides observed in the lower thermosphere typically originate in the middle and lower atmosphere then propagate upwards. As altitude increases, density decreases so in order to conserve momentum and energy, tidal amplitudes must increases with height. Tides generated below the mesopause will therefore dominate in the thermosphere over those generated in situ. This mechanism demonstrates how tides can play an important role in coupling between the lower- and upper- atmosphere.

Observations show that the semi-diurnal tide is dominant over the diurnal tide. This is because the semi-diurnal tide has stronger vertically propagating components than the diurnal. Diurnal components have smaller vertical wavelengths than those of the semidiurnal tide, meaning they are subject to destructive interference and preferential damping (*Hines* [1960], *Chapman and Lindzen.* [1970]). Tidal structure also varies with latitude as the amplitudes of vertically propagating modes decrease poleward of the equator. The propagating diurnal tide is found mainly between ±30°. Observations of high latitude oscillations are most likely due to ion-neutral interaction.

An analytical treatment of tidal oscillations in the atmosphere was presented by *Chapman and Lindzen* [1970], *Holton* [1975] and *Volland* [1988]. The theory uses tidal equations derived by *Laplace* in 1825 to calculate the atmosphere's response to global perturbations. The theory assumes that the atmosphere is in geostrophic balance i.e. the motion of the gas is governed by the pressure gradient and the Coriolis force alone. This assumption leads to simplified forms of the momentum and energy equations to which oscillations can be applied. The resulting differential equations are known as *Laplace's Tidal Equation*, which describes the global latitudinal structure of tidal amplitudes, and the *Vertical Structure Equation,* which describes the vertical structure of tides in the atmosphere. According to this theory, any global tidal perturbation can be decomposed into Eigenfunctions, typically referred to as *Hough modes*. These Hough modes are generally characterised by the longitudinal (zonal) wavenumber, s, and latitudinal (meridional) wave number n, written as (s,n). For

example, the diurnal tides have a longitudinal wave number of 1 meaning they are described by the family of *(1,n)* Hough modes. Similarly, the semidiurnal tides are described by the *(2,n)* modes. A standard notation is to use even values of n to describe modes symmetric to the geographical equator and odd values to describe asymmetric modes. Vertically propagating modes are described with positive values of n, while negative n describes non-propagating, or evanescent modes. Such oscillations do not propagate vertically and are confined to the region of forcing.

Tidal oscillations have a significant influence on the physical structure and energetics of the atmosphere. The variation in winds affects the dynamics and transport of chemical constituents. Associated changes in temperature and density can have a profound affect on composition. Non-linear interactions with other atmospheric oscillations such as planetary and gravity waves further enhance the influence tides have on the background atmosphere.

1.4.10 UPPER ATMOSPHERE PLANETARY WAVES

There is evidence that oscillations with periods typical of those associated with planetary waves (i.e. 2, 5 and 16 days) exist not only in the stratosphere and mesosphere, but also in the lower and even upper thermosphere (e.g. *Ward et al.* [1996], *Zhou et al.* [1997]). There are several mechanisms that may account for the presence of such oscillations in the thermosphere. These include in-situ planetary wave modulation of upward propagating gravity waves leading to an in-situ source of periodic forcing (*Meyer* [1999b]), planetary wave modulation of turbopause composition (through variations in temperature, vertical velocity and diffusion), and planetary wave modulation of upward propagating tides which result in planetary wave period variations of thermospheric tidal winds. The importance of each of these mechanisms is still not known and as such continues to be focus of much research.

1.5 ATMOSPHERIC NUMERICAL MODELS

By creating a plausible simulation of the atmosphere we can test the validity of scientific theories relating to the physical, radiative and chemical processes that control the real atmosphere. Sets of extreme conditions can be explored and, if the simulation

is sufficiently fast and sufficiently reliable, predictions can be made about the short and long term behaviour of the atmosphere. General circulation models or GCMs are models that solve the fundamental atmospheric equations in order to simulate the principal dynamical and energetic features of an atmosphere.

In order to reasonably represent the atmosphere with a numerical model, the basic equations that describe atmospheric structure and motion must be integrated with respect to time. Techniques such as finite differencing are employed such that changes in atmospheric behaviour can be simulated over time, following on from specified start up conditions. In 1918, *Lewis Fry Richardson* was the first to propose a method of predicting the evolution of tropospheric weather systems by solving finite difference expressions for the fundamental atmospheric equations. The first successful forecast was made in 1950 by *Charney*, *Fjörtoft* and *von Neumann* who assumed a barotropic atmosphere, i.e. a homogeneous atmosphere of uniform density with no vertical motion. Advances came with baroclinic models which allowed vertical motion but assumed quasi-geostrophy, i.e. no drag forces. Today, huge advances in computing capability mean we can use *Primitive equation models* which solve the basic equations of motion, continuity and energy and require a vast number of calculations to be done.

In order to solve the equations of atmospheric dynamics and momentum in three dimensions, two techniques are commonly employed. The first is the grid point method in which variable values, such as temperature or velocity, are defined at regularly spaced grid points in longitude, latitude and altitude. The equations are then solved by replacing the spatial derivatives with finite differences. The accuracy of such models will vary depending on the size of the time step and the resolution of the grid chosen. Such models require large amounts of computing resource, meaning the model solution usually reflects a balance between available computing power and the level of detail required. The second technique that is often used involves representing horizontal variations in terms of spherical harmonic expansions. In the vertical, the variables are evaluated at discrete grid points. These *spectral models* are valuable in that they require less computing time. The vertical coordinate system of models is

usually formulated in pressure or log pressure coordinates as will be discussed in section 2.5.

The first atmospheric GCMs concentrated on the troposphere as it was assumed that the rarefied upper atmosphere would have little influence on the terrestrial climate. Modern day GCMs are still largely split into two groups, those that deal with the lower and those that deal with the upper atmosphere. The U.K. Meteorological Office Unified Model, and the U.S. National Center for Atmospheric Research (NCAR) Coupled Community Model (CCM2) are two highly sophisticated lower atmosphere GCMs that play a central role in weather prediction and climate research. The first upper atmosphere GCM was developed in 1967 by *Kohl and King* and considered forces due to pressure gradients and ion drag alone. Pressure gradients were derived from semi-empirical density models, largely based on observations of satellite drag. Ion drag was derived from a globally uniform parameterisation of the ionosphere. Despite the simplicity of the representation, the model was able to reasonably reproduce upper thermospheric winds at mid and low latitudes.

The first global 3-dimensional, time-dependent GCMs were developed by *Fuller-Rowell and Rees* [1980] and *Dickinson* [1981]. These models self consistently solved the fully coupled non-linear time-dependent equations of energy and momentum in three dimensions. Since their development, these models have continually evolved. Realistic self-consistent calculations have been adopted over parameterisations of composition and ionospheric parameters (*Fuller-Rowell et al.* [1984], *Roble and Ridley* [1994], *Fuller-Rowell et al.* [1996a], *Millward et al.* [1996]). Dynamical processes such as tidal forcing have been introduced (*Müller-Wodarg* [1997]). The latest versions of these codes are regularly used in studies of the middle and upper atmosphere. The NCAR Thermosphere-Ionosphere-Mesosphere-Energetics (TIME) GCM (*Roble and Ridley* [1994]) extends from about 30km up to the upper thermosphere and is one of the principal time-dependent 3D GCMs used today. The model used in this thesis covers the same altitude range as TIME and is an extension of the code originally developed by *Fuller-Rowell et al.*, as described in *Harris* [2001].

It is worth mentioning that some lower atmosphere models do extend into the lower thermosphere, namely the United Kingdom Universities Global Atmospheric Modelling Program (UGAMP) extended GCMs, and Canadian Middle Atmosphere Model (CMAM). To date there are no 3D global GCMs that extend from ground level up and include the upper thermosphere-ionosphere system because until now, the development and application of such a complex model required prohibitively large amounts of computing resource. This requirement is being addressed with the ongoing development of models such as the Whole Atmosphere Community Climate Model (WACCM), which currently extends from the ground to the lower thermosphere.

Another class of models that are commonly used to simulate atmospheric behaviour are 2D or zonally averaged models. These require considerably less computing power than 3D models as they solve zonally averaged equations of energetics, dynamics and composition (e.g. *Crutzen* [1975], *Harwood and Pyle* [1975], *Garcia and Solomon* [1983], *Brasseur et al.* [1990], *Summers et al.* [1997], *Zhu et al.* [1997a]). 2D models are therefore well suited to long-term studies. One disadvantage arises from the difficulty associated with representing zonally asymmetric motions in zonally averaged equations. This is especially hard in the stratosphere where planetary waves cause a net meridional transport.

Models that solve for perturbations relative to a pre-defined background atmosphere are well suited to studies of planetary waves and tides (e.g. *Forbes* [1982a,b], *Hagan et al.* [1995]). The effects of wave-mean and wave-wave coupling can be investigated in great detail. However, only fully self-consistent models can represent the complex radiative-dynamic-chemical coupling mechanisms that go to define the structure and behaviour of the terrestrial atmosphere.

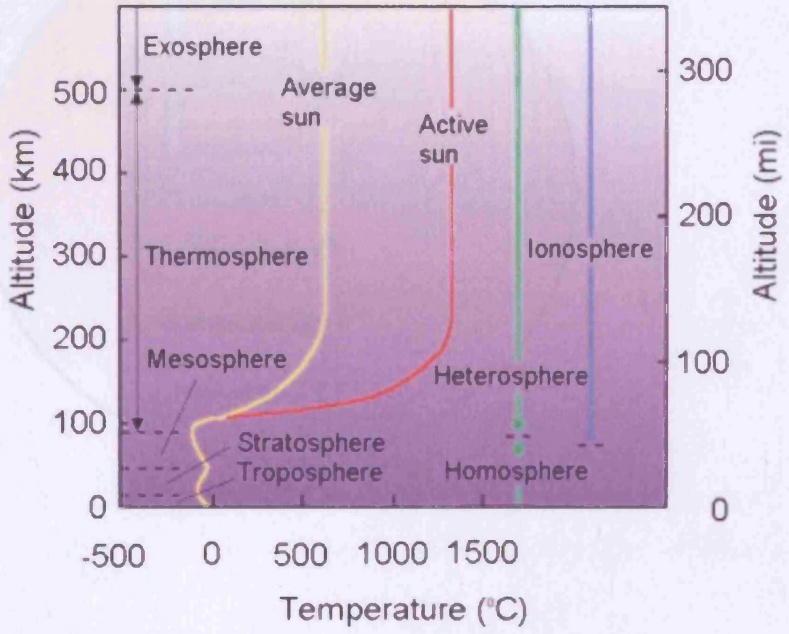


Figure 1.1 Temperature structure of the Terrestrial atmosphere

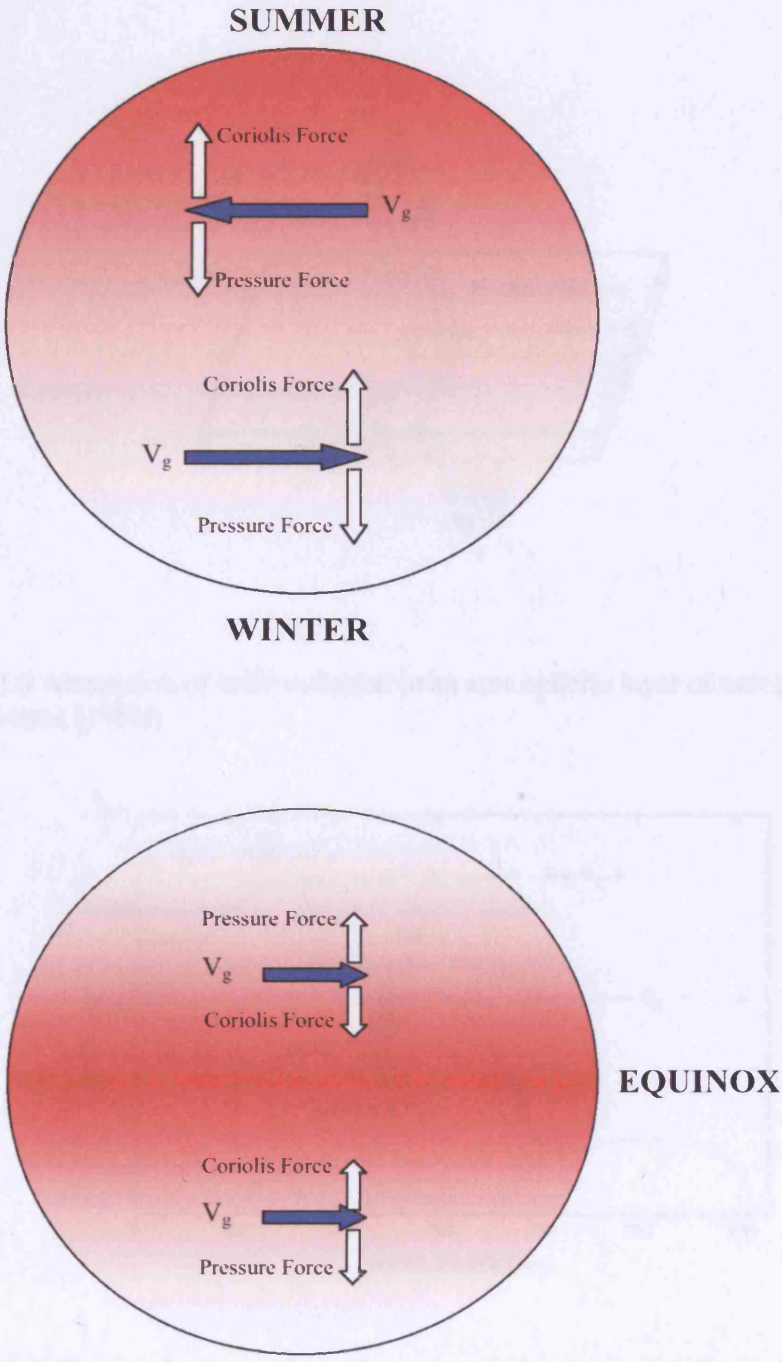


Figure 1.2 Diagram illustrating the mechanism for the mesospheric zonal wind jets (*Harris* [2001]).

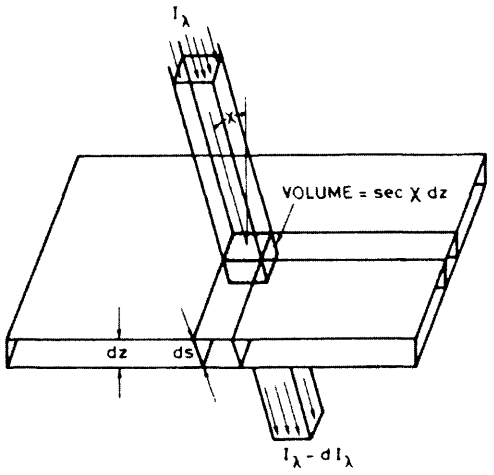


Figure 1.3 Absorption of solar radiation in an atmospheric layer of unit area (*Brasseur and Solomon* [1986]).

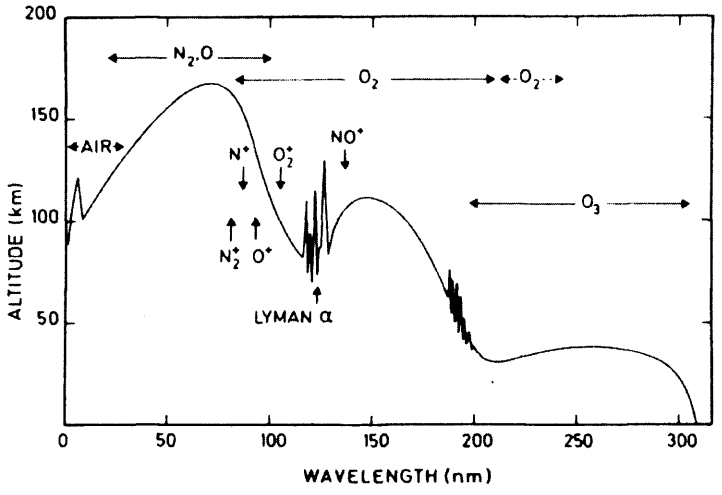


Figure 1.4 Depth of penetration of solar radiation as a function of wavelength. Altitudes correspond to an attenuation of 1/e. Principal absorbing species and ionisation limits are indicated (*Brasseur & Solomon* [1986]).

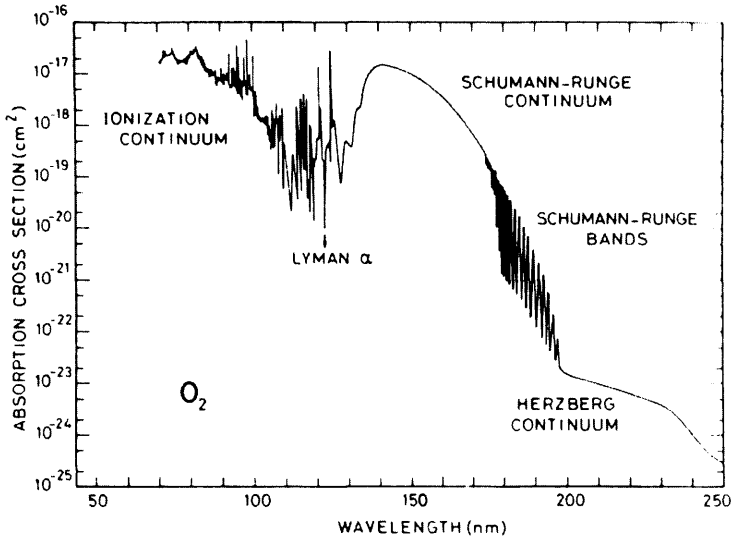


Figure 1.5 Spectral distribution of the absorption cross section of molecular oxygen. (*Brasseur and Solomon* [1986]).

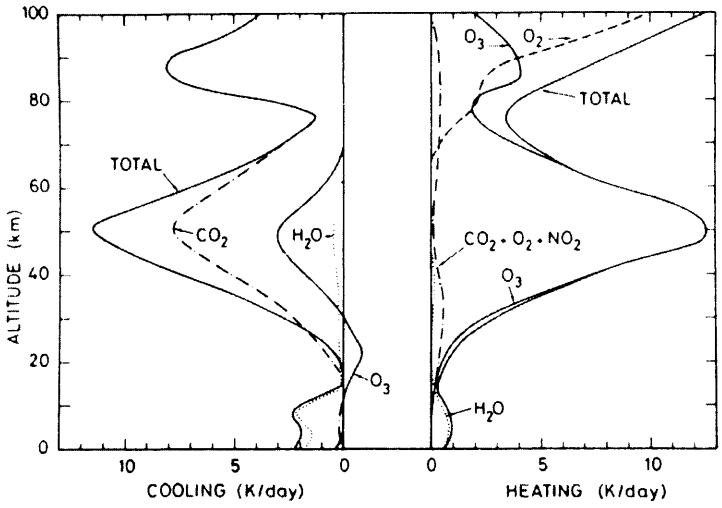


Figure 1.6 Vertical distribution of solar shortwave heating rates by O_3, O_2, NO_2, H_2O, CO_2, and of the terrestrial long wave cooling rates by CO_2, O_3, and H_2O. From *London* [1980].

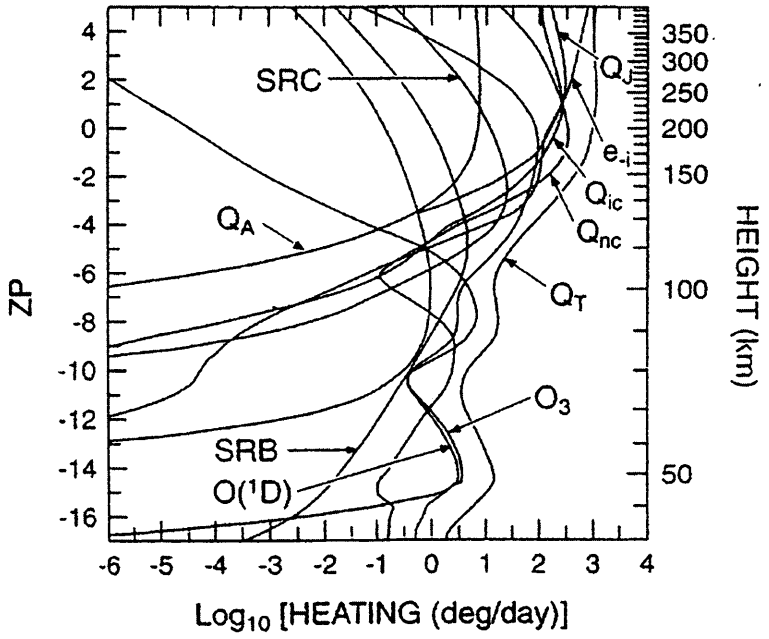


Figure 1.7 Global heating rates calculated by the NCAR globally averaged mesosphere thermosphere model for solar minimum (*Roble* [1995]). Units are $LOG_{10}(K\ day^{-1})$, where Q_T is the total heating rate; Q_{ic} heating due to ion-neutral exothermic chemical reactions; Q_{nc} heating due to neutral-neutral exothermic chemical reactions; Q_J Joule heating; Q_A heating due to particle precipitation; e_{-i} is heating due to collisions between thermal electrons, ions, and neutrals; O(^{1}D) heating due to quenching of O(^{1}D); SRC and SRC are heating due to absorption by O_2 in the Schumann-Runge continuum and bands; O_3 heating due to absorption by O_3 in the Hartley, Huggins, and Chappuis bands of ozone.

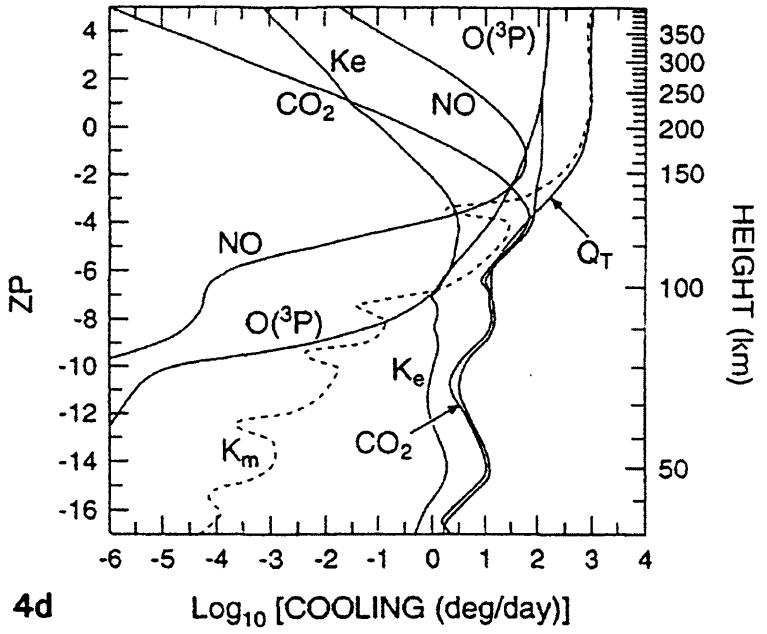


Figure 1.8 Global mean cooling rates calculated by the NCAR globally averaged mesosphere thermosphere model for solar minimum (*Roble* [1995]). Units are $LOG_{10}(K\ day^{-1})$, where Q_T is the total neutral gas cooling rate; K_m the cooling rate due to downward molecular thermal conduction; K_T the cooling rate due to eddy thermal conduction; NO radiative cooling due to 5.3μm emission from nitric oxide; CO_2 radiative cooling due to 15μm band emission of carbon dioxide; $O(^3P)$ radiative cooling due to 63μm fine structure emission of atomic oxygen; O_3 radiative cooling due to 9.6μm emission from ozone.

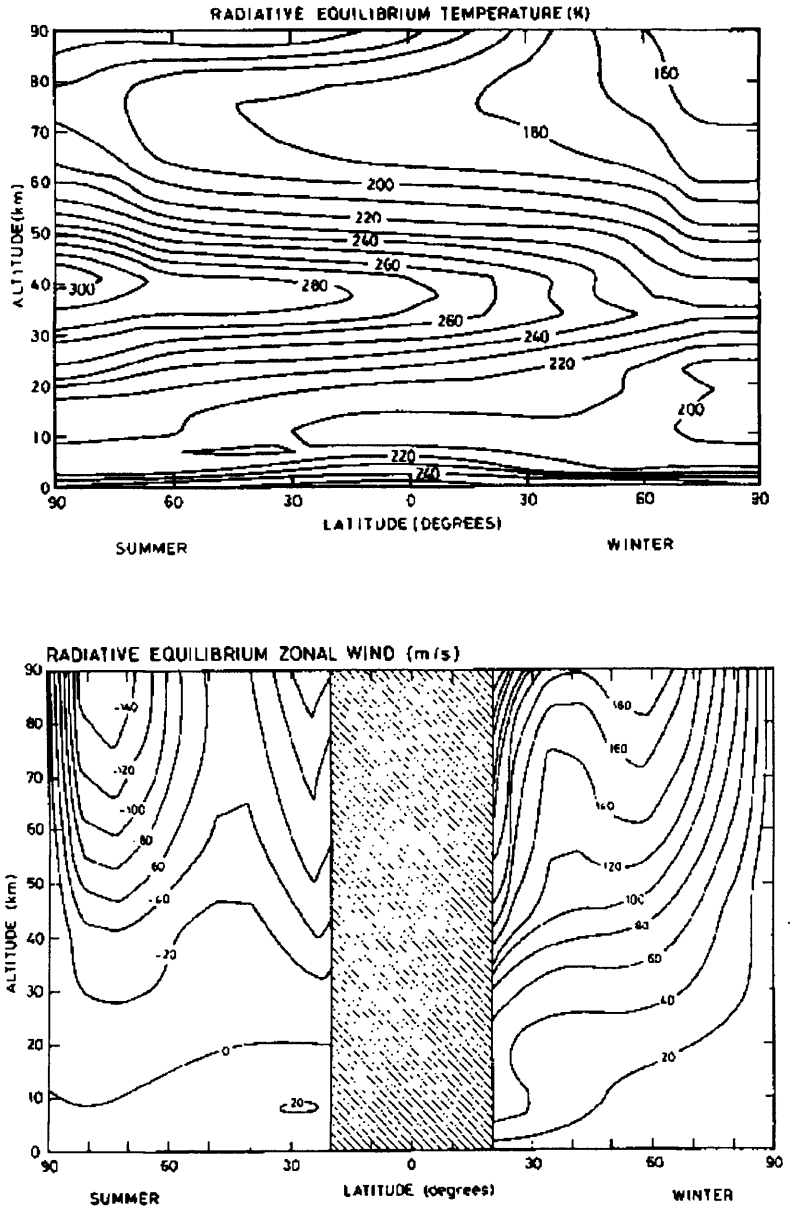


Figure 1.9 Calculated radiative equilibrium temperatures (*top*) and associated zonal winds (*bottom*). Positive values denote eastward winds and negative values denote westward winds. From *Geller* (1983).

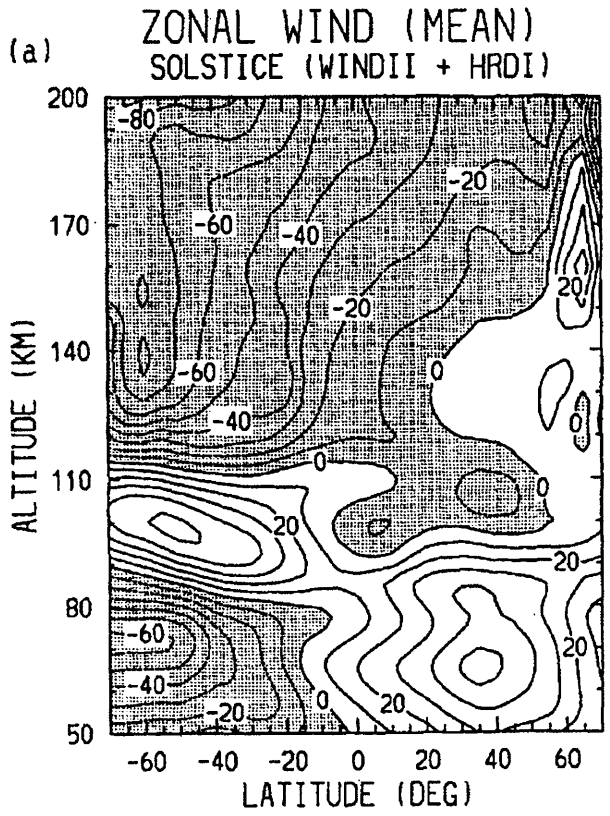


Figure 1.10 Mesosphere-thermosphere zonal mean zonal wind climatology as measured by the UARS satellite. Positive values denote eastward winds (*McLandress et al.* [1996a]).

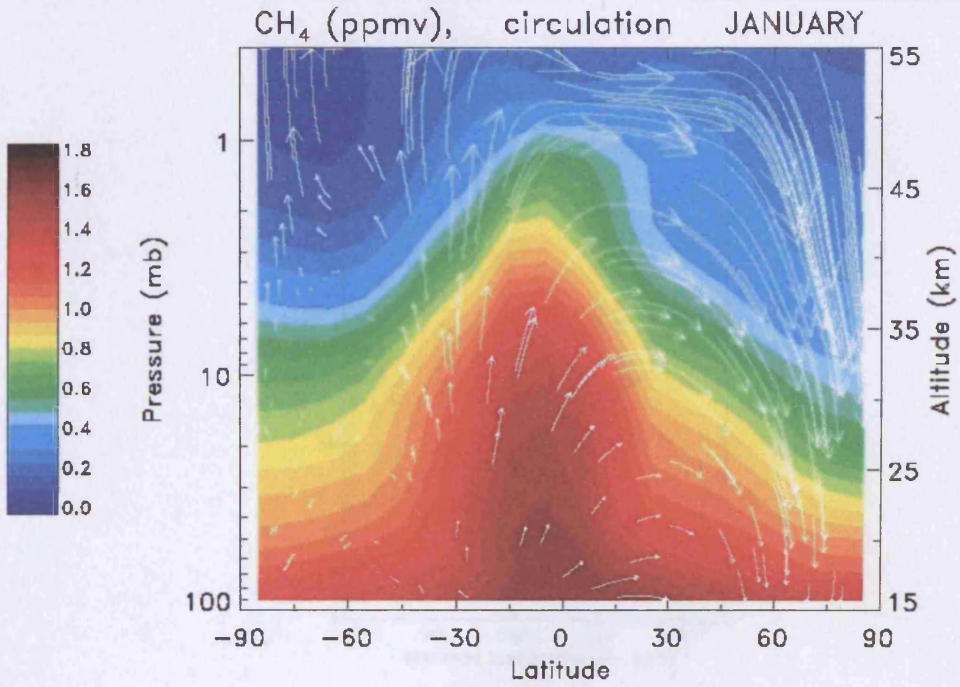


Figure 1.11 Zonally averaged methane volume mixing ratio (*ppmv*) in January. Superimposed is the Brewer-Dobson circulation pattern (arrows), *NASA* [2000].

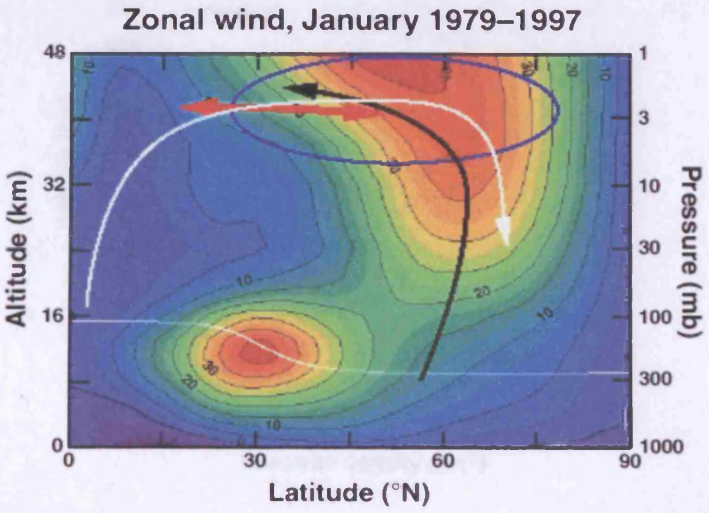


Figure 1.12 Zonal mean wind (positive eastward), planetary wave propagation (*black arrow*) and dissipation (*blue circle*), and resultant Brewer-Dobson circulation (*white* arrow), *NASA* [2000].

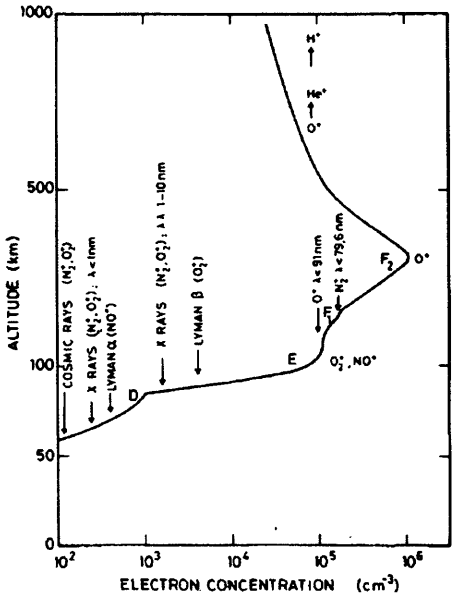


Figure 1.13 Definition of the ionospheric layers based on electron density distribution. Principal ionising sources are shown *Banks and Kockarts* [1973], taken from *Brasseur and Solomon* [1986].

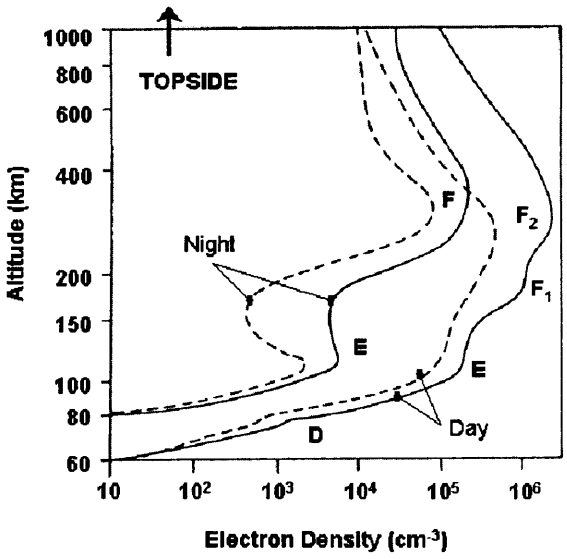


Figure 1.14 Diurnal variation of ionospheric layers during solar maximum (*solid line*) and solar minimum (*dotted line*). After *Hargreaves* [1995].

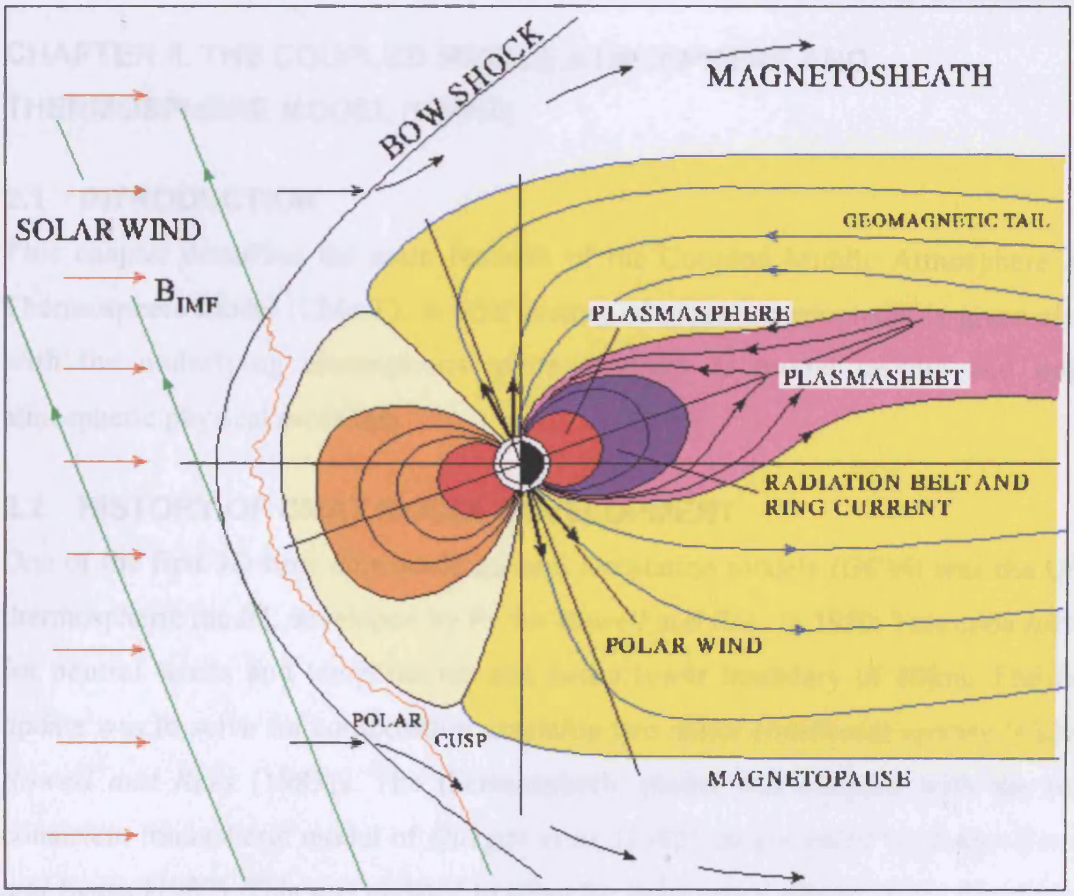


Figure 1.15 Schematic representation of the structure of the magnetosphere.

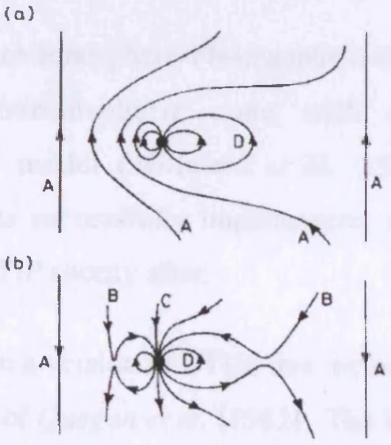


Figure 1.16 Illustration of closed and open magnetic field under conditions of (a) northward IMF (b) southward IMF, (*Hargreaves* [1979]).

CHAPTER II. THE COUPLED MIDDLE ATMOSPHERE AND THERMOSPHERE MODEL (CMAT)

2.1 INTRODUCTION

This chapter describes the main features of the Coupled Middle Atmosphere and Thermosphere model (CMAT). A brief history of model development is given along with the underlying assumptions made in order to model middle and upper atmospheric physical processes.

2.2 HISTORY OF CMAT MODEL DEVELOPMENT

One of the first 3D time dependent general circulation models (GCM) was the UCL thermospheric model, developed by *Fuller-Rowell and Rees* in 1980. This code solved for neutral winds and temperatures and had a lower boundary of 80km. The first update was to solve for composition assuming two major constituent species (*Fuller-Rowell and Rees* [1983]). The thermospheric model was coupled with the self-consistent ionospheric model of *Quegan et al.* [1982], as presented by *Fuller-Rowell and Evans* [1987]. This was updated to solve for three major species O, N_2, O_2 (*Fuller-Rowell et al.* [1984]). The resulting model is commonly known as the CTIM (Coupled Thermosphere-Ionosphere Model).

The Coupled Thermosphere-Ionosphere-Plasmasphere model (CTIP) was created by coupling the CTIM thermospheric code with a mid- and low-latitude ionosphere/plasmasphere model (*Millward et al.* [1996]). Self-consistent lower boundary tidal forcing was successfully implemented, primarily in CTIM (*Müller-Wodarg* [1997]), then in CTIP shortly after.

CMAT was developed from a version of CTIM that included tidal forcing and the high latitude ionospheric model of *Quegan et al.* [1982]. The lower boundary was extended downwards to 30km, vertical resolution was increased, and numerous changes were made to energetics, dynamics and composition as described in *Harris* [2001].

Other than the NCAR Thermosphere-Ionosphere-Mesosphere-Electrodynamics (TIME) GCM (*Roble and Ridley* [1994]), CMAT is the only 3D code that self consistently solves the fundamental atmospheric equations from 30km upwards, including the effects of the ionosphere. General circulation models such as CMAT can be used to study the response of the atmosphere to factors such as variable solar/geomagnetic activity, compositional changes and dynamical forcing. The extended lower boundary in CMAT makes it particularly useful for studies of the mesosphere/lower-thermosphere (MLT) region, thought to be a region of great importance for upper/lower atmospheric coupling.

2.3 THE BASIC ASSUMPTIONS

Certain assumptions must be made in order to solve the basic equations that collectively describe atmospheric dynamics, energetics and composition. Firstly, we assume a state of local thermodynamic equilibrium. This is acceptable over the height range of the model where collisional excitation/de-excitation processes dominate over radiative processes. Secondly, a state of hydrostatic equilibrium is assumed, i.e. the force due to gravity exactly balances the force due to the vertical pressure gradient. The atmosphere is treated as an ideal gas and the acceleration due to gravity g is treated as a constant.

2.4 THE FUNDAMENTAL EQUATIONS

The fundamental equations used within the model are based on three basic laws; the conservation of mass, the conservation of momentum (Newton's 2nd law), and the conservation of energy (1st law of thermodynamics).

2.4.1 THE CONTINUITY EQUATION

Consider a volume of dimensions $\Delta x \Delta y \Delta z$, through which gas flows along the x-axis. If u_1 and u_2 are the scalar velocities of the gas entering and leaving the volume, the mass flux of gas into and out of the volume are $\rho_1 u_1$ and $\rho_2 u_2$ respectively, where ρ_1 and ρ_2 are the mass densities of gas at the entrance and exit boundaries of the volume. The mass entering the volume in time Δt is $\rho_1 u_1 \Delta y \Delta z \Delta t$, while the mass leaving is

$\rho_2 u_2 \Delta y \Delta z \Delta t$. The mass accumulating is $\Delta\rho \Delta x \Delta y \Delta z = \rho_1 u_1 \Delta y \Delta z \Delta t - \rho_2 u_2 \Delta y \Delta z \Delta t$. Dividing both sides by Δt and the box volume $\Delta x \Delta y \Delta z$

$$\frac{\Delta\rho}{\Delta t} = -\left(\frac{u_2\rho_2 - u_1\rho_1}{\Delta x}\right) \qquad 2.1$$

For an infinitesimal volume, in three dimensions

$$\frac{\partial\rho}{\partial t} = -\overline{\nabla}_3 \overline{V}_3 \rho \qquad 2.2$$

This is the continuity equation, where $\overline{\nabla}_3$ is the three dimensional del operator, and $\overline{V}_3$ the three-dimensional velocity vector in the x-y-z coordinate system.

2.4.2 GEOPOTENTIAL

Geopotential $d\Phi$ is defined as the work required to raise an air parcel of unit mass through height dz in the Earth's gravitational field g, such that

$$d\Phi = gdz \qquad 2.3$$

2.4.3 THE MOMENTUM EQUATION

Rishbeth and Garriott [1969] derived the following expression to describe the conservation of momentum in the neutral atmosphere

$$\frac{d\overline{V}_3}{dt} = -2\overline{\Omega} \times \overline{V}_3 + \overline{g} - \frac{1}{\rho}\overline{\nabla}_3 P + \frac{1}{\rho}\overline{\nabla}_3(\mu\overline{\nabla}_3)\overline{V}_3 - \nu_{ni}(\overline{V}_3 - \overline{U}_{i3}) \qquad 2.4$$

From left to right, the terms on the right hand side represent coriolis, gravitational acceleration, pressure gradient, horizontal and vertical viscosity and ion drag. $\overline{\Omega}$ is the Earth's angular rotation vector, μ the sum of the molecular and turbulent coefficients of viscosity, ν_{ni} the neutral-ion collision frequency, $\overline{V}_3$ and $\overline{U}_{i3}$ the 3 dimensional neutral and ion wind velocity vectors. Vertical components can be ignored due to the assumption of hydrostatic equilibrium.

2.4.3.1 THE LAGRANGIAN DERIVATIVE

The horizontal coordinates used in CMAT are fixed relative to the Earth's surface, i.e. are fixed points in the atmosphere through which the fluid is seen to flow. The coordinate system is in 'Eulerian space'.

The left-hand side of the momentum equation (2.4) is the Lagrangian derivative, relating to the rate of change of velocity of a fluid parcel moving with the fluid (i.e. along a trajectory). The primitive equations must be converted from the Lagrangian to the Eulerian form. For any variable A, the following transformation can be applied.

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + (\overline{V}_3 \cdot \overline{\nabla}_3)A \qquad 2.5$$

The first term on the right hand side is the Eulerian derivative, representing the rate of change at a fixed point in space momentarily occupied by the fluid. The second term on the right hand is the non-linear advective term, where advection refers to the carrying of properties in a fluid. This accounts for the fact that a fixed point in space is occupied by successive parts of the moving fluid.

2.4.3.2 THE CORIOLIS TERM

If an object is moving with respect to a rotating reference frame, a Coriolis force arises which acts perpendicular to both the direction of motion and the axis of rotation. The force does no work but acts to change the direction of motion such that a body will be pushed in a clockwise direction in the northern hemisphere and in an anticlockwise direction in the southern hemisphere. The first right hand term of the momentum equation describes this apparent force.

The horizontal component of the Coriolis force is $f\overline{V} \wedge \hat{k}$ where $\overline{V}$ is the horizontal velocity vector and $f = 2\Omega\sin\phi$ is the Coriolis parameter, ϕ the latitude, $\hat{k}$ is the unit vector along the axis of rotation. The angular momentum Ω is made up of two components, the first is due to the Earth's rotation and the second due to the fact that

the atmosphere is a fluid constrained to move on a spherical surface. An eastward wind V_y at latitude ϕ will result in a total angular velocity of

$$2\Omega + \frac{V_y}{R\cos\phi} \qquad 2.6$$

where R is the radius of the Earth.

2.4.3.3 ACCELERATION DUE TO GRAVITY

The force due to gravity is taken as $9.5ms^{-2}$ at all heights and is the second term on the left hand side of the momentum equation.

2.4.3.4 THE PRESSURE GRADIENT TERM

Pressure gradients can be produced by direct heating such as solar radiation, or by dynamical propagation of fields. When a pressure gradient exists in a fluid, a force acts upon the fluid, causing it to move down the pressure gradient. The third term on the right hand side of equation 2.4 accounts for this force.

2.4.3.5 THE VISCOSITY TERM

Viscosity of a fluid acts to smooth velocity gradients in the wind field, particularly over short vertical distances. The fourth right hand term in equation 2.4 is the force due to viscous drag. Above the turbopause, molecular viscosity of the gas becomes increasingly dominant over the turbulent viscosity which plays an important role below the turbopause. Vertical wind shear tends to be greater than horizontal and as such the vertical component dominates (*Fuller-Rowell* [1981]).

2.4.3.6 THE ION DRAG TERM

The final term on the right hand side of the momentum equation is termed ion-drag and is due to collisions between ions and the neutral gas. Depending on the relative magnitudes of the ion drift velocity and the neutral wind velocity, each can be a driver or a drag on the other. If the ion drift velocity is larger than the neutral wind velocity as is the case at high latitudes, the ions can drive the neutral wind such that the ion drag term is a momentum source. At low latitudes where the convection electric field

is not driving the ions to high velocities, the ion-drag term acts to decelerate the neutral wind.

The momentum equation for ions can be written as (*Rishbeth* [1972])

$$\frac{d\overline{U}_{i3}}{dt} = \overline{g} + \overline{\nabla} \cdot \frac{P_i + P_e}{Nm_i} + \frac{e}{m_i}(\overline{E} + \overline{U}_{i3} \times \overline{B}) - \nu_{in}(\overline{U}_{i3} - \overline{V}_3) \qquad 2.7$$

where $\overline{U}_{i3}$ and $\overline{V}_3$ are the ion and neutral velocities, P_i and P_e are the ion- and electron partial pressures, N the ion or electron number density (assumed equal), e the electron charge, m_i the ion mass, $\overline{E}$ the electric field vector, $\overline{B}$ the magnetic field vector, and ν_{in} the ion (or electron)-neutral collision frequency. Note that the Coriolis term $(-2\Omega \times \overline{U}_{i3})$ can be neglected as ν_{in} is large. An analogous expression describes the motion of electrons such that in steady state conditions, the horizontal components of equation 2.7 for ions and electrons can be written

$$eN_i(\overline{E} + \overline{U}_i \times \overline{B}) - \rho_i \nu_{in}(\overline{U}_i - \overline{V}) = 0 \qquad 2.8$$

$$-eN_e(\overline{E} + \overline{U}_e \times \overline{B}) - \rho_e \nu_{en}(\overline{U}_e - \overline{V}) = 0 \qquad 2.9$$

where ρ_i and ρ_e are ion and electron mass densities, ν_{en} the electron-neutral collision frequency, and $\overline{U}_e$ the electron velocity.

Adding these two equations and assuming charge neutrality (eN_i=eN_e) gives

$$eN_i(\overline{U}_i \times \overline{B}) - eN_e(\overline{U}_e \times \overline{B}) - \rho_i \nu_{in}(\overline{U}_i - \overline{V}) - \rho_e \nu_{en}(\overline{U}_e - \overline{V}) = 0 \qquad 2.10$$

Substituting for current density $\overline{J} = eN_i\overline{U}_i - eN_e\overline{U}_e$ gives

$$J \times B - \rho_i \nu_{in}(\overline{U}_i - \overline{V}) - \rho_e \nu_{en}(\overline{U}_e - \overline{V}) = 0 \qquad 2.11$$

Rearranging we get

$$\nu_{in}(\overline{U}_i - \overline{V}) = \frac{1}{\rho_i}\overline{J} \times \overline{B} - \frac{\rho_e}{\rho_i}\nu_{en}(\overline{U}_e - \overline{V}) \qquad 2.12$$

By Newton's 3rd law,

$$\rho_i \nu_{in} = \rho_n \nu_{ni} \qquad 2.13$$

therefore

$$-\nu_{ni}(\overline{V} - \overline{U}_i) = \frac{1}{\rho_n}\overline{J} \times \overline{B} - \frac{\rho_e}{\rho_n}\nu_{en}(\overline{U}_e - \overline{V}) \qquad 2.14$$

The last right hand term may be ignored as $\rho_e << \rho_n$, thus removing the dependence on ion velocity. The resulting expression can be used to calculate ion drag.

In order to calculate the Current density *J*, a generalised version of Ohm's law can be used

$$\overline{J} = \sigma \cdot (\overline{E} + \overline{V} \times \overline{B}) \qquad 2.15$$

where σ is the conductivity tensor, $\overline{E}$ is an externally applied convection field or internal polarisation electric field, and $\overline{V} \times \overline{B}$ is the dynamo electric field. The conductivity tensor is a function of the magnetic dip angle, *I*, and the parallel-, σ_0, Pedersen-, σ_1 and Hall, σ_2, conductivities. *Pedersen* conductivity refers to that component in the direction of an applied electric field while *Hall* conductivity refers to the component perpendicular to the field. Both Pedersen and Hall conductivities are in the plane perpendicular to the magnetic field. Pedersen conductivity peaks in the E-region at about 130km while Hall conductivity has a peak at about 100km. The component of conductivity along the magnetic field due to a parallel electric field is described as the *longitudinal* or *parallel* conductivity and is dominant above altitudes of about 400km.

In the model we assume layer conductivity (*Rishbeth and Garriott* [1969]) such that if a layer is limited in vertical extent, any vertical currents produced by the dynamo will result in an accumulation of charge at the layer boundary. These 'polarisation charges' modify the $\overline{V} \times \overline{B}$ field until the resultant flow is horizontal. Using this assumption we

can reduce the 3x3 conductivity tensor to a 2x2 tensor, σ_l, representing layer conductivity

$$\sigma_l = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix} \qquad 2.16$$

where

$$\sigma_{xx} = \frac{\sigma_0\sigma_1}{\sigma_1\cos^2 I + \sigma_0 \sin^2 I} \approx \frac{\sigma_1}{\sin^2 I} \qquad 2.17$$

$$\sigma_{xy} = -\sigma_{yx} = \frac{\sigma_0\sigma_2 \sin I}{\sigma_1\cos^2 I + \sigma_0 \sin^2 I} \approx \frac{\sigma_2}{\sin I} \qquad 2.18$$

$$\sigma_{yy} = \frac{\sigma_1\sigma_0 \sin^2 I + (\sigma_1^2 + \sigma_2^2)\cos^2 I}{\sigma_1\cos^2 I + \sigma_0 \sin^2 I} \approx \sigma_1 \qquad 2.19$$

The Pedersen and Hall conductivities are calculated in the model using the expressions due to *Rishbeth and Garriott* [1969]

$$\sigma_1 = \frac{N_i e r}{B(1+r^2)} \qquad 2.20$$

$$\sigma_2 = \sigma_1 r \qquad 2.21$$

where r is the ratio of collision to gyro frequencies, given by $m_i \nu_{in}/(eB)$. m_i is the mean ion mass, e is electronic charge, N_i is electron or ion density, assumed equal. The magnetic field of the Earth, $\overline{B}$, is specified using empirical datasets, as is the electric field $\overline{E}$ (see sections 2.11.3 and 2.11.4). ν_{in} denotes the ion-neutral collision parameter, which for molecular ions NO^+ and O_2^+ is taken from *Schunk and Walker* [1973]

$$\nu_{in-molecular} = 4.34\times10^{-6}[N_2] + 4.28\times10^{-16}[O_2] + 2.44\times10^{-6}[O] \qquad 2.22$$

and for O^+ from *Salah* [1993]

$$\nu_{in-o+} = 6.82\times10^{-16}[N_2] + 6.66\times10^{-16}[O_2] + 3.42\times10^{-17}[O]\sqrt{T}(1.08 - 0.139\log_{10}T + 4.51\times10^{-3}(\log_{10}T)^2 \quad 2.23$$

2.4.4 THE ENERGY EQUATION

The first law of thermodynamics expresses the principle of conservation of energy. This states that the energy change of a system is equal to the net exchange of energy across its boundaries and for a compressible fluid, can be described by equation 1.6. This equation represents the rate of change of the internal energy density of a fluid. In order to calculate the temperature distribution in the atmosphere, we must derive an expression for the conservation of total energy density per unit mass. All sources and sinks of energy must be included, along with vertical and horizontal transport and conduction, and energy transformation to and from kinetic, internal and potential.

An expression for the kinetic energy density per unit mass can be found by multiplying the momentum equation (2.4) by $\overline{V}$. The rate of change of energy can be then determined by summing geopotential (as defined in equation 2.3), internal, and kinetic energy. A full description of the energy equation as it appears in CMAT will be given in section 2.6.7

2.5 THE COORDINATE SYSTEM

The equations presented so far have been in a Cartesian coordinate system where the x-axis is positive in a southward direction, the y-axis is positive eastwards and the z-axis points vertically upwards. A simplification of the mathematics can be achieved by using a spherical pressure coordinate system. Assuming hydrostatic equilibrium, the vertical coordinate z is easily transformed to a pressure coordinate P. Where necessary, the fundamental equations are also transformed from Lagrangian to Eulerian coordinates as described in section 2.4.3.1.

When transforming the basic equations into the 'P' coordinate system it is convenient to use a different notation for component vectors whereby $\overline{V}_2$ refers to the horizontal

components only. Three-dimensional vectors may be written $\overline{V_2} + \underline{k}\overline{V_z}$ where $\overline{V_z}$ is the vertical wind velocity in the 'z' coordinate system, and $\underline{k}$ is a unit vector along the z axis. Similarly, $\overline{\nabla}_z$ represents the two-dimensional del operator and the three dimensional operator becomes $\overline{\nabla}_z + \underline{k}\frac{\partial}{\partial z}$. In the '$P$' system the three-dimensional operator may be written $\overline{\nabla}_p + \underline{k}\frac{\partial}{\partial P}$.

For any scalar S, the transformation between z and P coordinate systems in the x direction is given by

$$\left(\frac{\partial S}{\partial x}\right)_p = \left(\frac{\partial S}{\partial x}\right)_z + \left(\frac{\partial z}{\partial x}\right)_p \left(\frac{\partial S}{\partial z}\right) \qquad 2.24$$

From equation 2.3 for geopotential, we have

$$\left(\frac{\partial z}{\partial x}\right)_p = \frac{1}{g}\left(\frac{\partial \Phi}{\partial x}\right)_p \qquad 2.25$$

Substituting this and the hydrostatic equilibrium equation (1.1) into equation 2.24 gives

$$\left(\frac{\partial S}{\partial x}\right)_z = \left(\frac{\partial S}{\partial x}\right)_p + \rho\left(\frac{\partial \Phi}{\partial x}\right)_p \left(\frac{\partial S}{\partial P}\right) \qquad 2.26$$

Similarly in the y and t directions

$$\left(\frac{\partial S}{\partial y}\right)_z = \left(\frac{\partial S}{\partial y}\right)_p + \rho\left(\frac{\partial \Phi}{\partial y}\right)_p \left(\frac{\partial S}{\partial P}\right) \qquad 2.27$$

$$\left(\frac{\partial S}{\partial t}\right)_z = \left(\frac{\partial S}{\partial t}\right)_p + \rho\left(\frac{\partial \Phi}{\partial t}\right)_p \left(\frac{\partial S}{\partial P}\right) \qquad 2.28$$

The horizontal del operator in the z system must also be transformed

$$\overline{\nabla}_z S = \overline{\nabla}_p S + \rho \, \overline{\nabla}_p \Phi \frac{\partial S}{\partial P} \qquad 2.29$$

Similarly for a two dimensional vector $\overline{A}_2$

$$\overline{\nabla}_z \overline{A}_2 = \overline{\nabla}_p \overline{A}_2 + \rho \overline{\nabla}_p \Phi \frac{\partial \overline{A}_2}{\partial P} \qquad 2.30$$

In the vertical direction, using the equation of hydrostatic equilibrium

$$\frac{\partial S}{\partial z} = \frac{\partial S}{\partial P}\frac{\partial P}{\partial z} = -g\rho \, \frac{\partial S}{\partial P} \qquad 2.31$$

$$\frac{\partial \overline{A}_2}{\partial z} = \frac{\partial \overline{A}_2}{\partial P}\frac{\partial P}{\partial z} = -g\rho \frac{\partial \overline{A}_2}{\partial P} \qquad 2.32$$

2.6 TRANFORMED EQUATIONS

For a full derivation of the transformed equations listed below see *Fuller-Rowell and Rees* [1980], and *Fuller-Rowell* [1981].

2.6.1 THE HYDROSTATIC EQUATION

Substituting Φ for S in equation 2.31, the vertical transform becomes

$$\frac{\partial \Phi}{\partial P} = -\frac{1}{\rho} \qquad 2.33$$

This is the pressure coordinate version of the hydrostatic equation.

2.6.2 PRESSURE GRADIENT

The horizontal pressure gradient $\overline{\nabla}_z P$ can be expressed in pressure coordinates by using equation 2.29 and noting that $\overline{\nabla}_p P = 0$.

$$\overline{\nabla}_z P = \rho \overline{\nabla}_p \Phi \qquad 2.34$$

2.6.3 VERTICAL VELOCITY

Horizontal velocity is defined the same way in both the 'z' and 'P' coordinate systems. The vertical velocity however is different and in the pressure coordinate system is defined as

$$w = \frac{dP}{dt} \qquad 2.35$$

where w represents the vertical velocity relative to the isobaric surface. Positive/negative w indicates descending/ascending air relative to the isobaric surface.

The relationship between V_z and w can be found by applying the Eulerian expansion of the total derivative $\frac{d}{dt}$ in the pressure coordinate system, to the geopotential Φ

$$\frac{d\Phi}{dt} = \left(\frac{\partial\Phi}{\partial t}\right)_p + \overline{V}_2.\overline{\nabla}_p\Phi + w\frac{\partial\Phi}{\partial P} \qquad 2.36$$

Applying equations 2.3 and 2.33 gives

$$\frac{d\Phi}{dt} = g\frac{dz}{dt} = \left(\frac{\partial\Phi}{\partial t}\right)_p + \overline{V}_2\overline{\nabla}_p\Phi - \frac{w}{\rho} \qquad 2.37$$

therefore

$$V_z = \frac{dz}{dt} = \frac{1}{g}\left(\frac{\partial\Phi}{\partial t}\right)_p + \frac{1}{g}\overline{V}_2\overline{\nabla}_p\Phi - \frac{w}{\rho g} \qquad 2.38$$

The sum of the first two right hand terms in equation 2.38 is the *barometric vertical velocity* and refers to vertical motion in the z coordinate system of an isobaric surface due to expansion and contraction. The final component on the right hand side is the *divergence vertical velocity*, which describes the vertical velocity relative to an isobaric layer due to convergences and divergences in horizontal velocity.

2.6.4 THE PERFECT GAS LAW

$$P = \frac{RT\rho}{M} = gH\rho \qquad 2.39$$

where R is the universal gas constant.

2.6.5 THE CONTINUITY EQUATION

$$\frac{\partial w}{\partial P} + \overline{\nabla}_p \overline{V} = 0 \qquad 2.40$$

2.6.6 THE MOMENTUM EQUATION

In the pressure coordinate system the momentum equation is expressed as

$$\frac{\partial \overline{V}_2}{\partial t} = -\overline{V}_2 . \overline{\nabla}_p \overline{V}_2 - w\left(\frac{\partial \overline{V}_2}{\partial P}\right) - \overline{\nabla}_p \Phi - \left(2\Omega + \frac{V_y}{R\sin\theta}\right)\cos\theta \underline{k} \times \overline{V}_2 + g\frac{\partial}{\partial P}\left((\mu_m + \mu_t)\frac{P}{H}\frac{\partial \overline{V}_2}{\partial P}\right) + \frac{\mu_m}{\rho}\overline{\nabla}_p^2 \overline{V}_2 + \frac{1}{\rho}\left(\overline{J} \times \overline{B}\right) \qquad 2.41$$

where R is the radius of the Earth, w vertical velocity, Ω the rotation rate of the Earth, P pressure, μ_m and μ_t the coefficients of molecular and turbulent viscosity, $\overline{J}$ the current density and $\overline{B}$ the magnetic field.

The first 2 terms on the right hand side represent horizontal and vertical advection, the third is geopotential, the fourth Coriolis, the fifth vertical viscous drag, the sixth horizontal viscous drag and the last ion drag as discussed in section 2.4.3.6.

When converted to spherical coordinates and split into meridional and zonal components where V_θ is southward and V_ϕ eastward velocity we have

$$\frac{\partial V_\theta}{\partial t} = -\frac{V_\theta}{R}\frac{\partial V_\theta}{\partial \theta} - \frac{V_\Phi}{RSin\theta}\frac{\partial V_\theta}{\partial \phi} - \omega\frac{\partial V_\theta}{\partial P} - \frac{g}{R}\frac{\partial h}{\partial \theta} + \left(2\Omega + \frac{V_\phi}{R\sin\theta}\right)V_\phi Cos\theta$$
$$+ g\frac{\partial}{\partial P}\left[(\mu_m + \mu_t)\frac{p}{H}\frac{\partial V_\theta}{\partial p}\right] + \frac{1}{\rho}\left(\overline{\nabla}_p^2 V_\theta + \frac{1}{R^2}\frac{\partial \mu}{\partial \theta}\frac{\partial V_\theta}{\partial \theta} + \frac{1}{R^2\sin^2\theta}\frac{\partial \mu}{\partial \phi}\frac{\partial V_\theta}{\partial \phi}\right) \qquad 2.42$$
$$+ \frac{1}{\rho}\left(\overline{J} \times \overline{B}\right)$$

$$\frac{\partial V_\phi}{\partial t} = -\frac{V_\theta}{R}\frac{\partial V_\phi}{\partial \theta} - \frac{V_\phi}{RSin\theta}\frac{\partial V_\phi}{\partial \Phi} - \omega\frac{\partial V_\phi}{\partial p} - \frac{g}{RSin\theta}\frac{\partial h}{\partial \phi} - \left(2\Omega + \frac{V_\phi}{R\sin\theta}\right)V_\theta Cos\theta$$
$$+ g\frac{\partial}{\partial p}\left[(\mu_m + \mu_t)\frac{p}{H}\frac{\partial V_\phi}{\partial p}\right] + \frac{1}{\rho}\left(\overline{\nabla}_p^2 V_\phi + \frac{1}{R^2}\frac{\partial \mu}{\partial \theta}\frac{\partial V_\phi}{\partial \theta} + \frac{1}{R^2\sin^2\theta}\frac{\partial \mu}{\partial \phi}\frac{\partial V_\phi}{\partial \phi}\right) \qquad 2.43$$
$$+ \frac{1}{\rho}\left(\overline{J} \times \overline{B}\right)$$

where θ is co-latitude and ϕ longitude. For a derivation of equations 2.41, 2.42 and 2.43 see *Fuller-Rowell and Rees* [1980], *Fuller-Rowell* [1981] and *Müller-Wodarg* [2000].

2.6.7 THE ENERGY EQUATION

Expressions for the rate of change of internal energy of a gas, the kinetic energy density and the potential energy per unit mass are combined, then converted into the pressure coordinate system. The result is the energy equation as it appears in CMAT

$$\frac{\partial \in}{\partial t} + \overline{V}_2 \cdot \overline{\nabla}_p(\in + \Phi) + w\frac{\partial(\in + \Phi)}{\partial P} = Q_{EUV} + Q_{IR} + g\frac{\partial}{\partial P}\frac{(K_M + K_T)}{H}P\frac{\partial T}{\partial P} +$$
$$\frac{1}{\rho}(K_M + K_T)\overline{\nabla}_p^2 T - \qquad 2.44$$
$$g\frac{\partial}{\partial P}\frac{K_T g}{C_p} + \frac{\overline{J} \cdot \overline{E}}{\rho} + \overline{V}_2 \cdot g\frac{\partial}{\partial P}\frac{\mu}{H}P\frac{\partial}{\partial P}\overline{V}_2$$

The function $\in = 1/2(\overline{V}_2^2) + C_p T$ is defined as the specific enthalpy plus kinetic energy density per unit mass. The second and third terms on the left hand side represent horizontal and vertical advection.

The right-hand side gives individual terms for solar heating, infrared cooling, vertical molecular and turbulent conduction, horizontal molecular and turbulent conduction, vertical turbulent heat conduction due to the adiabatic lapse rate, ion drag and joule heating, and vertical viscous drag. K_M and K_T are coefficients of molecular and turbulent conductivity. A full derivation can be found in *Fuller-Rowell* [1981].

2.7 MAJOR SPECIES TRANSPORT

The three major species in the model are atomic and molecular oxygen, and molecular nitrogen. Major composition is calculated for each constituent using the continuity equation in the form

$$\frac{\partial \chi_i}{\partial t} = \frac{1}{\rho} m_i S_i - V_2 \cdot \overline{\nabla}_p \chi_i - w \frac{\partial}{\partial P} \chi_i - \frac{1}{\rho} \overline{\nabla}_p \cdot (n_i m_i C_i) + \frac{1}{\rho} \overline{\nabla}_p (D_{EDDY} n \overline{\nabla}_p m \chi_i) \qquad 2.45$$

(*Fuller-Rowell* [1984]), where $\chi_i = n_i m_i / \rho$ are mass mixing ratios of species i (i= O, O_2, N_2), S_i denotes chemical sources and sinks, m_i molecular mass of species i and m the mean molecular mass, n_i number density, C_i molecular diffusion velocity, and D_{EDDY} the eddy diffusion coefficient. The first right hand term corresponds to chemical production and loss, the second and third horizontal and vertical advection, and the fourth and fifth molecular and turbulent diffusion respectively.

The diffusion velocity, C_i, for each major constituent is found by simultaneously solving the generalised diffusion equation for a multiple-species, non-uniform gas (*Chapman and Cowling* [1952])

$$\frac{1}{n} \sum_{i \neq j} \left(\frac{\chi_j}{m_i D_{ij}} n_j m_j C_j - \frac{\chi_i}{m_j D_{ij}} n_i m_i C_i \right) = \overline{\nabla} \chi_i + \frac{\chi_i}{m} \overline{\nabla} m + \left(1 - \frac{m_i}{m} \right) \frac{\chi_i}{P} \overline{\nabla} P \qquad 2.46$$

where D_{ij} is the mutual molecular diffusion coefficient of species i through species j. The mean molecular mass m is given by

$$m = \frac{n_1 m_1 + n_2 m_2 + n_3 m_3}{n} \qquad 2.47$$

where $n=n_1 + n_2 + n_3$ is the total number density. A weighted mean of the velocity through each constituent is then taken.

2.7.1 MOLECULAR DIFFUSION COEFFICIENTS

Density gradients in the atmosphere cause gas molecules to flow in a direction opposite to the density gradient. The flux of molecules is proportional to the density gradient ($\Phi = -D\nabla n$), where the proportionality factor, D, is the molecular diffusion coefficient.

The molecular diffusion coefficients used in the model are assumed symmetric, $D_{ij} = D_{ji}$, and are shown in Table 2.1 below

i-j	D_{ij} [$m^2 s^{-1}$]
$O - O_2$	$[T/273.0]^{1.75} \cdot [\ 1.01325 \cdot 10^5/P\ ] \cdot 2.60 \cdot 10^{-5}$
$O - N_2$	$[T/273.0]^{1.75} \cdot [\ 1.01325 \cdot 10^5/P\ ] \cdot 2.60 \cdot 10^{-5}$
$O_2 - N_2$	$[T/273.0]^{1.75} \cdot [\ 1.01325 \cdot 10^5/P\ ] \cdot 1.81 \cdot 10^{-5}$

Table 2.1 Model mutual molecular diffusion coefficients, after (*Colgrove* [1966]), where P is pressure in pascals and T is temperature in kelvin.

2.7.2 EDDY DIFFUSION COEFFICIENT

The turbulent diffusion coefficient profile, D, is taken from *Roble* [1995]. Attempts have been made to calculate the diffusion coefficient self-consistently in the gravity wave drag parameterisation but this was found to give unpredictable results and unrealistic wind velocity profiles (*Harris* [2001]).

2.8 MINOR SPECIES TRANSPORT

The minor constituent composition is calculated in a similar way to major composition except the transport effects of minor constituents back onto the major constituents are

neglected. This means that molecular diffusion is not considered mutual and is analogous to turbulent diffusion such that equation 2.45 becomes.

$$\frac{\partial \chi_i}{\partial t} = \frac{1}{\rho} m_i S_i - V_2 \cdot \overline{\nabla}_p \chi_i - w \frac{\partial}{\partial P} \chi_i + \frac{1}{\rho} \overline{\nabla}_p ((D_{EDDY} + D_{MOLEC}) n \overline{\nabla}_p m \chi_i) \qquad 2.48$$

2.8.1 MINOR SPECIES DIFFUSION COEFFICIENTS

The molecular diffusion coefficients of N(^{4}S), N(^{2}D) and NO through O, O_2 and N_2 are due to *Levin et al.* [1990]. The empirical expressions of *Banks and Kockarts* [1973] are used for diffusion of all other minor species through N_2 and O_2, whereas diffusion through O is calculated using the empirical expression of *Fuller et al.* [1966].

The overall diffusion coefficient of a minor species through the major constituents is calculated using the weighted means of the diffusion coefficients through each constituent (*Wilke* [1950])

$$D_{1-mixture} = \frac{1}{v_2 / D_{1-2} + v_3 / D_{1-3} + ...} \qquad 2.49$$

where v_n is the volume mixing ratio of the major constituent, and D_{1-n} is the molecular diffusion coefficient of the species through that major constituent.

The concept of chemical families will be discussed in section 2.14.2 but it is worth noting here that the values of D_{1-n} for chemical families are weighted means of the diffusion coefficients of each family member through the major constituent under consideration.

2.9 THE CMAT MODEL GRID

The CMAT model grid points are fixed with respect to the surface of the Earth and are spaced every 18° longitude, 2° latitude, and 1/3 scale height vertically. 63 pressure levels span a height range from approximately 30km (10mb) to between 250 and 600km (7.586×10^{-9} mb) depending on solar activity.

From equation 1.3, if the scale height H, is the height over which pressure P, decreases by a factor e.

$$P = P_0 e^{-z/H} = P_0 e^{\Delta_n (1-n)} \qquad 2.50$$

where Δ_n=1/3, and n is the dimensionless pressure coordinate extending from 1 to 63.

In order to maintain numerical stability, the vertical resolution of the levels between two and one scale height below the upper boundary are one scale height instead of 1/3. The reasons for this shall be discussed in the next section.

2.10 NUMERICAL INTEGRATION METHOD

The fundamental equations described require a numerical solution to solve the non-linear terms. A finite difference integration scheme is used to generate solutions at each grid point. The technique requires that first-order horizontal and vertical spatial derivatives be expressed in their two-point centred-difference approximation given by

$$\frac{dF_j}{dx} = \frac{F_{j+1} - F_{j-1}}{2\Delta_j} \qquad 2.51$$

where $F_j = F(x_j)$ is any function, Δ_j is the step length and x_j is a point on a grid such that

$$x_j = x_1 + \sum_{n=1}^{j-1} \Delta x_n \qquad 2.52$$

The second order derivative is given by

$$\frac{d^2F_j}{dx^2} = \frac{F_{j+1} - 2F_j + F_{j-1}}{\Delta_j^2} \qquad 2.53$$

where the subscripts denote the points on an x-axis at j-1, j , j+1, an equal distance Δ_j apart.

At any time step, the value of a particular parameter is evaluated by using its value at time t-1, then integrating the differential expressions defined above over the length of the step. This is an iterative process, and is repeated for all grid points.

In order to integrate expressions such as the energy and momentum equations which take the form $\frac{\partial F(t)}{\partial t} = LF(t)$, a finite approximation with respect to time is applied. Ignoring terms higher than 2nd order, integrating over time Δt, the solution becomes

$$F_{n+1} = F_n + LF_n \Delta t \qquad 2.54$$

where $F(t)$ is any function and L is the spatial differential operator containing the first and second order derivatives. This is the Euler method and is stable for differential equations that contain 2nd order spatial derivatives, but not for equations containing 1st order spatial derivatives. A numerical diffusion term is added in order to increase stability of those equations that contain first order terms, such as advective terms, details of which can be found in *Harris* [2001]. The numerical scheme used in the model is referred to as an explicit Forward Time Centred Space (FTCS) scheme.

Note that the major and minor constituent advection scheme in CMAT is not 'centred difference' but 'upstream difference'. This scheme is better suited to representing sharp discontinuities in constituent concentrations caused by fast photochemistry in the lower thermosphere and mesosphere. The upsteam differencing scheme is numerically diffusive so an antidiffusion correction is applied (*Smolarkiewicz* [1983]). Details of this scheme and associated correction can be found in *Harris* [2001].

2.10.1 NUMERICAL INSTABILITIES

The modified Euler method is stable for differential equations that contain second order spatial derivatives. An example of this would be diffusion type equations,

$$\frac{\partial F}{\partial t} + K\frac{\partial^2 F}{\partial x^2} = 0 \qquad 2.55$$

for integration time steps of $\Delta t < 0.5\Delta^2/K$ where K is a diffusion coefficient, and Δ is the grid step length in the x direction.

In the upper thermosphere where molecular diffusion is the dominant vertical transport process, the time step Δt required to maintain numerical stability becomes very small (<10s). To run CMAT with a 10s time step over the whole vertical range would substantially increase model run time and as such be impractical. To overcome this problem a vertical resolution of 1 scale height is used for the top two model pressure levels, meaning an integration time step of 60s can be safely used over the whole vertical range of the model.

2.11 MODEL PARAMETERS

2.11.1 THE COEFFICIENTS OF THERMAL CONDUCTIVITY

The coefficient of molecular heat conduction, K_m, is obtained from a simple empirical expression (*Rees* [1989])

$$K_m = K_0 T^s \qquad [\mathrm{Wkg^{-1}K^{-1}m^{-1}}] \qquad 2.56$$

where the numerical parameters K_0 and s are given for each of the major constituents. A number density weighted mean due to O, O_2, and N_2, is taken to find the overall value of K_m.

The coefficient of turbulent heat conductivity, K_t, is obtained by applying the relationship $K_t=c_p\rho D$, where D is the turbulent or eddy diffusion coefficient ($\mathrm{m^2s^{-1}}$), ρ is the mass density and c_p the specific heat capacity at constant pressure.

2.11.2 THE COEFFICIENTS OF VISCOSITY

The coefficient of molecular viscosity μ_m is calculated from an expression by *Dalgarno and Smith* [1962], such that

$$\mu_m = 4.5 \cdot 10^{-5} \left(\frac{T}{1000} \right)^{0.71} \qquad [\mathrm{kg\ m^{-1}\ s^{-1}}] \qquad 2.57$$

The coefficient for turbulent viscosity is calculated from the coefficient of turbulent heat conductivity using,

$$\mu_t = \frac{2K_t}{c_p} \text{ [kg m}^{-1}\text{ s}^{-1}\text{]} \qquad 2.58$$

where c_p is specific heat capacity at constant pressure, and the factor 2 is the Prandtl number used within the model, being the ratio of kinematic viscosity to thermal diffusivity.

2.11.3 MAGNETIC FIELD

The strength of the magnetic field in the model is set using the International Geomagnetic Magnetic Field (IGRF, *see Langel* [1992]). This is an empirical representation of the Earth's magnetic field based on worldwide geomagnetic measurements, and includes the offset dipole nature of the geomagnetic field.

2.11.4 ELECTRIC FIELD AND PARTICLE PRECIPITATION MODELS

As the solar wind plasma blows across the Earth's magnetic field lines, a horizontal convective electric field is set up. This field is mapped along the conducting magnetic field lines onto the high latitude ionosphere. This field drives currents which give rise to Joule heating, an important energy source in the thermosphere. High-energy particles precipitating into the auroral zones also impart energy to the thermosphere.

The high latitude electric field model used in CMAT is based on the statistical model of *Foster et al.* [1986] in which the mean electric field is deduced from measurements of plasma motion within the F-region using the Millstone Hill incoherent scatter radar. *Codrescu et al.* [2000] showed that the observed small-scale electric field variability and the mean electric field contribute equally to thermospheric global Joule heating. To account for this, a random value taken from a Gaussian distribution of standard deviation 0.01Vm^{-1} is added to the existing mean field x and y components, when they exceed 0.001Vm^{-1}.

The Foster electric field model is used in conjunction with a precipitation model based on TIROS/NOAA satellite auroral particle measurements (see *Fuller-Rowell and Evans* [1987]). In this model, the total energy flux carried towards the atmosphere by electrons and positive ions, integrated over the energy range 300eV to 100keV, is used

to create statistical maps of energy influx. The shape of the energy spectrum is determined by the 'characteristic energy', which would be the average energy of the distribution for a Maxwellian particle spectrum. CMAT is provided with a specific auroral activity level (the TIROS level) related to the geomagnetic activity index as given in Table 2.2. The level of activity specified is used to determine the electric field patterns, the amount of auroral energy input to the model and thus the ion production rates for the major species due to particle precipitation.

Codrescu et al. [1997] showed that medium energy particles in the range 30keV to 2.5MeV had a significant effect on electron density and composition in the polar mesosphere. This energy range is accounted for in CMAT through use of the precipitation patterns developed by *Codrescu et al.*

Average K_p	1^-	1^o	1^+	2^-	2^+	3^-	3^+	4^o	5^-	6^-
Tiros Level	1	2	3	4	5	6	7	8	9	10

Table 2.2 Tiros Precipitation/K_p index relationship

2.11.5 THE HIGH LATITUDE IONOSPHERE MODEL

The high latitude ionosphere model used in CMAT calculates electron density, electron and ion temperatures, field aligned velocities and distribution of ion species assuming photochemical equilibrium and is described in detail in *Quegan et al.* [1982] and *Fuller-Rowell and Evans* [1987].

2.12 ENERGETICS

2.12.1 SOLAR IRRADIANCE

The absorption of solar radiation in CMAT is determined using the Lambert-Beer exponential absorption law (equation 1.27) as described in section 1.3.7. Within the CMAT model there is an upper boundary. The contribution to the integral in equation 1.27 due to the atmosphere above that boundary must be accounted for. This is done

using the expression of *Rishbeth and Garriott* [1969], who showed that if a species has local number density n_0, and local scale height H_0, the total number density per unit area, or *column density*, above that point is given by n_0H_0.

Photoionisation and dissociation rates are calculated using the expressions given in section 1.3.8.

2.12.2 SOLAR HEATING

The heating rate due to absorption of solar radiation at wavelength λ is given by

$$Q(\lambda) = n\frac{hc}{\lambda}\varepsilon(\lambda)\sigma_a(\lambda)I_\infty(\lambda)\exp\left(-\tau(\lambda)\right) \qquad 2.59$$

where n is the constituent concentration, hc/λ is the photon energy, and h is Planck's constant. $\varepsilon(\lambda)$ is the heat efficiency that represents the amount of energy which is converted to heat, σ_a is the absorption cross section of the constituent, $I_\infty(\lambda)$ is the intensity of radiation at wavelength λ at 1AU and $\tau(\lambda)$ is the optical depth as defined in equation 1.28.

The total heat production over a specified wavelength interval is obtained by integrating equation 2.59 over the wavelength range.

2.12.3 THE THERMOSPHERIC HEATING ROUTINE

The CMAT thermospheric heating scheme, developed by Fuller-Rowell, calculates heating in the UV and EUV regions at 5nm spectral intervals and includes the effects of individual line emissions. The following expressions due to *Swinder and Gardener* [1967] are used to estimate the Chapman grazing incidence function for each absorbing species.

$$Ch(x,\chi \le \frac{\pi}{2}) = \left(\frac{\pi x}{2}\right)^{\frac{1}{2}}\left(1 - erf\left(x^{\frac{1}{2}}\cos\frac{\chi}{2}\right)\right)\exp\left(x\cos^2(\frac{\chi}{2})\right)$$
$$Ch(x,\chi \ge \frac{\pi}{2}) = \left(\frac{\pi x}{2}\sin\chi\right)^{\frac{1}{2}}\left(1 + erf(-\cot\chi\left(\frac{x\sin\chi}{2}\right)^{\frac{1}{2}}\right)\left(1 + \frac{3}{8x\sin\chi}\right) \quad 2.60$$

where $x=(a+z)/H$, a being the radius of the Earth, z altitude, and H the scale height of the absorber being considered. This accounts for the curvature of the Earth at large solar zenith angles.

EUV fluxes within the spectral range 1.8-105nm are modified *Hinteregger* [1970] fluxes due to *Torr et al.* [1979]. The Schumann-Runge continuum UV fluxes between 105-180nm are due to *Torr et al.* [1980c].

The absorption and ionisation cross sections of O, O_2 and N_2 between 1.8-105nm are due to *Torr et al.* [1979]. Cross sections for wavelengths below 5nm are taken from *Oshio et al.* [1966] and *Banks and Kockarts* [1973]. O_2 Schumann-Runge cross sections are taken from *Torr et al.* [1980a].

EUV heating efficiencies due to *Roble et al.* [1987] are applied in order to account for the channelling of photon energy into processes such as secondary electron events, electron collisional processes, and airglow loss which are not currently solved for in the model. These heating efficiencies also account for the contribution due to ion-neutral exothermic reactions.

Solar heating in the UV spectral range is primarily due to O_2 dissociative heating between 130-240nm. In the lower thermosphere, the lifetime of O is greater than a day and the energy absorbed during O_2 photolysis can be stored as chemical energy. This energy is lost when the O and its associated chemical potential energy are transported from the dissociation site (100-130km). Below 100km the energy is released through recombination reactions. UV heating efficiencies are not applied in CMAT, instead the

O-O bond energy is subtracted and the later recombination exothermic heating subsequent to transport is calculated self consistently.

Dissociative heating due to absorption by O_2 of radiation in the Schumann-Runge bands (175-200nm) is an important heating mechanism in the lower thermosphere. This is calculated using the following parameterisation due to *Strobel* [1978]

$$\frac{Q_{SRB}}{[O_2]} = \frac{1}{0.67[C_{O2}] + 3.44 \times 10^9 [C_{O2}]^{\frac{1}{2}}} \qquad 2.61$$

where $[O_2]$ is local number density in cm^{-3}, $[C_{O2}]$ total slant column density of O_2 in cm^{-2}, and Q_{SRB} the heating rate in 0.1 Wm^{-3}. If $[C_{O2}] < 10^{18}$ cm^2, then

$$\frac{Q_{SRB}}{[O_2]} = 2.43 \times 10^{-19} \qquad 2.62$$

2.12.4 THE MESOSPHERIC HEATING ROUTINE

The calculation for mesospheric solar heating includes the principal sources due to dissociation of O_3 in the Hartley (242-310nm) and Huggins (310-400nm) bands. Absorption by O_2 of radiation in the Hertzberg continuum (200-242nm), along with a contribution due to dissociation of O_3 in the Chappuis (400-850nm) band is also calculated. Fluxes and absorption cross sections from *Strobel* [1978] are used to calculate the Hartley continuum dissociation rate coefficient. Heating is calculated using the following heating algorithms due to *Strobel* [1978]

$$Q_{HERZ} = 1.5 \times 10^3 (6.6 \times 10^{-24} O_2 + 4.9 \times 10^{-18} O_3) \times \exp(-6.6 \times 10^{-24} C_{O2} - 4.9 \times 10^{-18} C_{O3}) \qquad 2.63$$

$$\frac{Q_{HART}}{C_{O3}} = 4.8 \times 10^{-14} \times \exp(-8.8 \times 10^{-18} C_{O3}) \qquad 2.64$$

$$\frac{Q_{HU}}{O_3} = \frac{1}{C_{O3}}\begin{bmatrix}4.66\times10^{3} - 7.8\times10^{2}\exp(-1.77\times10^{-19}C_{O3})\\ -3.88\times10^{3}\exp(-4.22\times10^{-18}C_{O3})\end{bmatrix} \quad 2.65$$

$$\frac{Q_c}{O_3} = 1.05\times10^{-15}\exp(-2.85\times10^{-21}C_{O3}) \quad 2.66$$

where Q_{HERZ} is the Hertzberg heating rate in Wm^{-3}, which contains a contribution from O_2 absorption. Q_{HART}, Q_{HU} and Q_C are heating rates due to absorption in the Hartley, Huggins and Chappuis bands in 0.1 Wm^{-3}. O_3 is local number density cm^{-3}, C_{O3} total slant column density of O_3.

Solar cycle variation is based on *Brasseur* [1993] and is 3% for the Hartley, Huggins and Chappuis bands, and 4% for the Hertzberg continuum.

In the dense stratosphere, solar radiation can undergo multiple scattering whereby a particle abstracts energy from the incident beam and reradiates it in all directions. The optical depth describing the attenuation of radiation is thus the result of absorption *and* scattering. CMAT does not account for scattering, which can act to increase the heating rates in the upper stratosphere by about 20% between 30-40km for a surface albedo of 0.3 (*Brasseur and Solomon* [1986]).

2.12.5 MESOSPHERIC HEATING EFFICIENCIES

Heating efficiencies in the model account for a 5-20% loss in heating efficiency due to airglow loss mechanisms, as calculated by *Mlynczak and Solomon* [1993].

2.12.6 ATOMIC OXYGEN COOLING

Radiative cooling above about 110km is dominated by the fine structure 63μm emission of atomic oxygen. To account for this, CMAT uses the *Bates* [1951] parameterisation

$$Q_{IR(O)} = \frac{1.67\cdot10^{-25}\exp(-228.0/T_n)}{1+0.6\exp(-228.0/T_n)+0.2\exp(-325.0/T_n)}\cdot[O] \quad 2.67$$

where T_n is neutral temperature, and [O] is atomic oxygen number density (m^{-3}). Due to uncertainty in validity (see discussion in *Harris* [2001]), an overall reduction factor of 0.5 is applied, after *Roble et al.* [1987]. Overestimation of cooling in the lower thermosphere, due to increasing opacity below 120km, has lead to the following damping factor being applied (*Fuller-Rowell*, private communication [1998]).

$$Q_{IR(O)-NEW} = Q_{IR(O)}\left(\exp\frac{-[Z-80.0]}{6.4}\right) \qquad 2.68$$

where Z is altitude in km. Below 80km the atomic oxygen cooling is set to zero.

2.12.7 NO RADIATIVE COOLING

5.3 μm non-LTE band emission of NO is an important source of cooling between about 150-200km. This is calculated using the following expression from *Kockarts* [1980]

$$Q_{NO} = h\upsilon n\omega A_{10}\left(\frac{g_1}{g_0}\right)\exp\left(\frac{-h\upsilon}{kT}\right) \qquad 2.69$$

where Q_{no} is the cooling rate (ergs s^{-1}), n the number of NO molecules in the vibrational level v=0, and $h\upsilon$ is the energy of the photon emitted in the spontaneous emission process (1876 cm^{-1})

$$NO_{v=1} \rightarrow NO_{v=0} + h\upsilon$$

A_{10} is the transition probability, taken as 13.3 s^{-1} after *King and Crawford* [1972], g_1 and g_0 are the statistical weights for the levels v=1 and v=0, assumed equal. k is the Boltzmann constant and T the kinetic temperature. ω is a dilution factor, representing a non-LTE departure from a Boltzmann distribution of energy levels, given by

$$\omega = \frac{k_{10}n(M)}{k_{10}n(M) + A_{10}} \qquad 2.70$$

where k_{10} is the deactivation rate coefficient of NO by collisions with constituent M between vibrational levels v=1 and v=0. n is the concentration of constituent M. Only deactivation by atomic oxygen is considered since $k_{10}(N_2) < k_{10}(O_2) << k_{10}(O)$, where $k_{10}(O)$ is 6.5×10^{-11} $cm^3 s^{-1}$ at 300K after *Fernando and Smith* [1979].

2.12.8 CO_2 COOLING

Cooling due to Local Thermodynamic Equilibrium (LTE) and non-LTE 15μm band emission of CO_2 is the primary cooling mechanism in the upper mesosphere. CMAT uses the parameterisation of *Formichev et al.* [1998] as outlined in *Harris* [2001]. The following collisional deactivation rate constants (k) in $cm^3 s^{-1}$ of $CO_2(01^10)$ by O_2, N_2, and O are used (*Shved et al.* [1998])

$$k_{N_2} = 5.5 \times 10^{-17} \sqrt{T} + 6.7 \times 10^{-10} \exp(-83.8 T^{-\frac{1}{3}}) \quad 2.71$$

$$k_{O_2} = 10^{-15} \exp(23.37 - 230.9 T^{-\frac{1}{3}} - 564 T^{-\frac{2}{3}}) \quad 2.72$$

$$k_O = 3 \times 10^{-12} \quad 2.73$$

Deactivation by collision with atomic oxygen is the dominant mechanism within the non-LTE region. There is still widespread uncertainty about the value of the associated rate constant, which has a significant impact on calculations of the energy budget in the MLT region.

2.12.9 O_3 COOLING

9.6μm band emission of ozone plays a minor role in cooling between 20-75km. CMAT uses the parameterisation of *Formichev and Shved* [1985] to account for this.

2.12.10 EXOTHERMIC CHEMICAL HEATING

As previously mentioned, above about 100km ultraviolet radiation acts to dissociate molecular oxygen, leading to the formation of atomic oxygen. At altitudes above 80km the lifetime of atomic oxygen is greater than a day and the chemical potential energy created by the photolysis reaction can be transported away from the reaction site.

CMAT calculates the release of energy due to neutral-neutral reactions following the transport of constituents. Heating efficiencies must be applied to certain exothermic reactions to account for energy losses from radiatively active products, such as the vibrationally active hydroxyl radical created in the reaction between hydrogen and ozone

$$H + O_3 \rightarrow OH(v=1\text{-}9) + O_2$$

This reaction is a major contributor to exothermic heating in the middle atmosphere but the magnitude of the loss due to the decay of the excited OH radical is still uncertain. A heating efficiency of 0.6 is applied in CMAT after *Mlynczak and Solomon* [1993].

2.13 DYNAMICS

2.13.1 LOWER BOUNDARY FORCING

Daily zonal means of geopotential height, as given by the Mass Spectrometer and ground based Incoherent Scatter Extended MSIS-E90 empirical model (*Hedin* [1991]), are applied at the lower boundary of CMAT. The date dependence of the data enables representation of a realistic seasonally varying wind structure.

The lower boundary zonal mean temperature and wind fields in CMAT can be set by several methods as described in *Harris* [2001] but are generally evaluated by linearly extrapolating values from the two over lying levels.

2.13.2 TIDAL FORCING

Tidal forcing is applied at the lower boundary of CMAT using self-consistent geopotential height perturbations as derived from classical tidal theory (*Müller-Wodarg* [1997]). An outline of their implementation in CMAT can be found in *Harris* [2001]. The (1,1), (2,2), (2,3), (2,4) and (2,5) Hough modes, as described in section 1.4.9, are implemented in the model. These Hough modes describe the horizontal structure of tidal oscillations. For each mode a phase shift corresponding to the local time of the maximum is specified, along with the tidal amplitude. Phases and amplitudes at the CMAT lower boundary are taken from Hough decompositions of the

Global Scale Wave Model (*M. Hagan*, private communication, [1999]), a table of which can be found in *Harris* [2001].

The lower boundary wind and temperature perturbations associated with tidal oscillations are calculated as described in the previous section, following the addition of the tidal and MSIS-E90 geopotential height fields.

2.13.3 THE GRAVITY WAVE PARAMETERISATION

The gravity wave parameterisation used in CMAT is due to *Meyer* [1999a,b] and is described as a hybrid *Lindzen-Matsuno* scheme, referring to the schemes of *Lindzen* [1981] and *Matsuno* [1982]. The Lindzen scheme considers a spectrum of discrete waves defined at a launch level. The vertical propagation of each individual wave is considered conservative until the wave perturbation produces a superadiabatic lapse rate and the wave begins to break. At this point, the eddy mixing induced by the dissipating wave is assumed sufficient to prevent any further amplitude growth. The following expression due to *Lindzen* [1981] describes the eddy diffusion K_{GW} that satisfies this condition

$$K_{GW} = Ak\frac{(\overline{u}-c_i)^3}{N^3}\left(\frac{1}{2H}-\frac{3}{2}\frac{d\overline{u}/dz}{(\overline{u}-c_i)}\right) \qquad 2.74$$

where A is a scaling factor accounting for wave intermittency, k is the characteristic horizontal wavenumber, $\overline{u}$ the background velocity, N the Brunt–Väisälä frequency, c_i the horizontal phase speed of the i^{th} wave, H the scale height and z the vertical coordinate.

The vertical flux of horizontal momentum $F_{GW\text{-}LZ}$ associated with the wave breaking is given by

$$F_{GW-LZ} = -\frac{N^2 K_{GW}}{(\overline{u}-c_i)} \qquad 2.75$$

The *Matsuno* [1982] scheme assumes that viscosity (molecular and eddy) is present everywhere in the atmosphere and that viscous dissipation of gravity waves causes momentum flux deposition

$$F_{GW-M} = -\frac{F(c,\tilde{u})}{\rho}\frac{\partial \tau}{\partial z} \qquad 2.76$$

where ρ is density, $F(c,\tilde{u})$ is the momentum flux at the gravity wave source level, $\tilde{u}$ is the wave amplitude, related to momentum flux $u'w'$ at the source level by $\tilde{u}$ =(u'w'2N/k)$^{1/3}$ after *Holton* [1982]. τ is the transmissivity of the wave between two height levels, z_1 and z_2, and is given by

$$\tau(z_1,z_2) = \exp\left[-\int_{z_2}^{z_1} \frac{2K_{tot}(z')N^3}{k(c_i - u(z'))^4}dz'\right] \qquad 2.77$$

where K_{tot} is the diffusion coefficient due to both molecular and turbulent diffusion. The value of τ approaches zero as the viscosity increases and the wave is totally absorbed.

Meyer [1999a,b] proposed the following expression for the altitude of the 'breaking level' where waves begin to dissipate

$$z_b = z_0 + 3H\ln\left[\frac{c_i - u}{\tau^{2/3}\tilde{u}}\right] \qquad 2.78$$

It should be noted that molecular diffusion is not included in a pure Lindzen scheme and its inclusion in the transmissivity calculation has an impact on the calculated breaking heights.

In order to calculate the gravity wave drag terms in CMAT, the critical level above which a wave is assumed to have been totally absorbed is first calculated for each wave in the spectrum. This level is the upper limit for calculations of gravity wave momentum deposition and is set at the point where the phase speed of the wave is

equal to the background wind, or where $|u-c_i|<2ms^{-1}$. The transmissivity of each wave is then calculated using the combined turbulent and molecular diffusion coefficient. Currently a 1D global profile of eddy diffusion due to *Roble* [1995] is used in this calculation, not a self-consistent value as suggested in *Meyer* [1999a,b]. The breaking heights z_b, are calculated at each model level up to the critical level using expression 2.78 above. If the height of a level in question is greater than the breaking height but less than the critical level then wave dissipation is assumed. Note that a wave can intermittently break as it propagates upwards and the combined wave-mean lapse rates change. The Lindzen turbulent diffusion coefficients K_{GW} are then calculated and summed over the whole wave spectrum for each model height. These are used in equation 2.75 to calculate the flux of horizontal momentum $F_{GW\text{-}LZ}$ due to *Lindzen* wave dissipation at each breaking height. To avoid a sharp transition in calculated gravity wave drag at the point where the first wave starts to break, an exponential decay is applied to $F_{GW\text{-}LZ}$ below the breaking height. This process is repeated in the zonal and meridional directions to give terms that are added to the momentum equation.

The implementation of this parameterisation in *Meyer* [1999a,b] is slightly different to that described above in that wave transmissivity is calculated using the *Lindzen* eddy diffusion coefficient K_{GW} due to gravity wave breaking. An iterative process is performed whereby once the transmissivity and breaking heights have been calculated, K_{GW} is recalculated and fed back into the transmissivity and breaking height calculations until a point is reached where the breaking heights converge. As such, the breaking height and flux calculations for one wave take into account the eddy diffusion produced by other waves. Currently the calculation of breaking height for one wave in CMAT is not influenced by the behaviour of the rest of the spectrum. It was found that the CMAT peak drag values and wind profiles compared better with other models when the scheme was implemented in this way.

Another difference between the scheme in CMAT and *Meyer* [1999a,b] is that Meyer used the combined Matsuno and Lindzen momentum fluxes, $F_{GW\text{-}LZ}$ and $F_{GW\text{-}M}$, whereas CMAT only uses $F_{GW\text{-}LZ}$ in the momentum equation. $F_{GW\text{-}LZ}$ dominates in the mesosphere (50-90km) where the only significant contribution of the Matsuno scheme

is in the calculation of breaking heights. However $F_{GW\text{-}M}$ is important in the lower thermosphere where the influence of molecular diffusion (and/or viscosity) is increasingly important. When developing the CMAT scheme, it was assumed that the contribution of $F_{GW\text{-}M}$ to the total wave drag was insignificant.

2.13.4 THE GRAVITY WAVE SPECTRUM

The gravity wave spectrum used in the *Meyer* [1999a,b] scheme in CMAT is a Gaussian distribution consisting of 19 waves in both the zonal and meridional directions. The maximum horizontal phase velocity is set to 60ms^{-1}. A full description can be found in *Harris* [2001].

2.13.5 PLANETARY WAVE DRAG

Planetary wave dissipation in the upper winter stratosphere plays an important role in circulation and associated transport and variability of chemical constituents. Realistic planetary wave forcing of the lower boundary of CMAT is not implemented as initial simulations were unable to reproduce realistic wave dissipation. In order to approximate planetary wave effects, *Harris* [2001] implemented zonal and meridional Rayleigh friction terms that are tuned to be consistent with the planetary wave drag morphology profiles of *Brasseur et al.* [1990]. This is a crude approximation that will be resolved with the release of a variable resolution version of the model, currently under development.

2.14 COMPOSITION

The neutral chemical scheme in CMAT is based on the mesospheric oxygen and hydrogen scheme of *Allen et al.* [1984], CO and CO_2 chemistry of *Solomon et al.* [1985], and mesospheric odd nitrogen reactions of *Garcia and Solomon* [1983], as outlined in *Fuller-Rowell* [1993]. JPL-97 reaction rate coefficients from *DeMore et al.* [1997] are applied. The distribution of D region ion species is calculated assuming photochemical equilibrium using the incorporated Sheffield University Coupled High Latitude Ionosphere Model (*Quegan et al.* [1982], *Fuller-Rowell et al.* [1996a]).

2.14.1 PHOTOELECTRON PRODUCTION, IONISATION AND DISSOCIATION

As mentioned in section 1.3.8 primary photon ionisation rates of O, O_2 and N_2 are calculated using expression 1.30. Dissociation rates of O_2 and N_2 are determined using expression 1.32. Ionisation by high energy photons can result in photoelectrons that have sufficient energy to ionise other particles. The possible products of ionisation or dissociation by a primary photon and subsequent reactions by energetic primary and secondary photoelectrons are given below (*Fuller-Rowell* [1993])

$$O_2 + \begin{matrix} h\nu \\ e^* \end{matrix} \rightarrow O_2^+ + \begin{matrix} e^* \\ 2e^* \end{matrix}$$

$$\rightarrow O^+ + O + \begin{matrix} e^* \\ 2e^* \end{matrix}$$

$$\rightarrow O + O + (e^*)$$

$$O + \begin{matrix} h\nu \\ e^* \end{matrix} \rightarrow O^+ + \begin{matrix} e^* \\ 2e^* \end{matrix}$$

$$N_2 + \begin{matrix} h\nu \\ e^* \end{matrix} \rightarrow N_2^+ + \begin{matrix} e^* \\ 2e^* \end{matrix}$$

$$\rightarrow N^+ + N + \begin{matrix} e^* \\ 2e^* \end{matrix}$$

$$\rightarrow N + N + (e^*)$$

The relative partitioning between the possible products of primary photon ionisation depends on the cross sections for each process. For altitudes above 100km in CMAT, height dependant ratios based on the model of *Strickland and Meier* [1982] determine the fraction of each ion produced by the primary photon ionisation e.g. ionisation of N_2 is partitioned between N^+ and N_2^+. The tables of fractional production of N^+, N_2^+, $O^+$and O_2^+ used in CMAT are given in *Fuller-Rowell* [1993]. Below 100km, the ratios are kept constant to the lower boundary.

For wavelengths below 35nm, the photoelectrons produced during the major species ionisation process are sufficiently energetic that they can cause one or more further

ionisations. Vertical profiles of the ratio of ionisation due to photoelectrons, compared to the primary ionisation rate of N_2 are applied in CMAT. Above 100km altitude, these ratios are based on photoelectron distributions calculated by the model of *Strickland and Meier* [1982] for an electron energy spectrum of up to 450eV. Below this altitude, values are estimated from the available energy of photons that would penetrate to the given altitude, as given in *Fuller-Rowell* [1993]. Similar ratios are applied for O and O_2 above 100km, with extrapolated values below. It is worth noting that all ratios used are the mean of two profiles from the *Strickland and Meier* model, one for low and one for high solar activity, and as such show no variation with solar cycle.

2.14.1.1 NIGHT-TIME IONISATION

Night-time ionisation of NO occurs due to solar Lyman alpha radiation that is scattered off the hydrogen geocorona. The night-time ionisation rate in CMAT is set to be 1000 smaller than the daytime rate for an overhead sun (*Brasseur and Solomon* [1986]). Night-time ionisation rates of O, O_2 and N_2 are set with the same factor, accounting for ionisation from galactic EUV emission.

2.14.1.2 HIGH LATITUDE IONISATION

As discussed in section 2.11.4, CMAT uses a particle precipitation model based on TIROS data to simulate high latitude energy inputs. The total ionisation due to precipitating high-energy particles is calculated for each TIROS activity level. The ionisation rates of each major constituent are then calculated using the following branching ratios due to *Jones and Rees* [1973].

$$Q_{N_2} = \frac{0.92[N_2]}{\Delta} Q_T \qquad 2.79$$

$$Q_{O_2} = \frac{[O_2]}{\Delta} Q_T \qquad 2.80$$

$$Q_O \frac{0.56[O]}{\Delta} Q_T + 0.5 Q_{O_2} \qquad 2.81$$

where

$$\Delta = 0.92[N_2] + 1.5[O_2] + 0.56[O] \qquad 2.82$$

In the equations above, Q_n is the total ionisation rate of constituent n in s^{-1}, and Q_T is the total ionisation rate due to particle precipitation. $[n]$ is the number density of constituent n. The numerical factors in these equations take into account differences in the magnitudes of the ionisation cross sections.

2.14.2 CHEMICAL FAMILIES AND PARTITIONING

In section 1.3.6 the concept of chemical lifetimes was discussed. When using a time dependent numerical model to calculate the variation in time of a species that has a very short chemical lifetime, one must use a time step that is shorter than the lifetime. This can be computationally expensive and leads to the introduction of chemical families whose lifetimes can be much longer than those of their constituent members. Take for example the *odd oxygen family*, O_x, made up of $O(^3P)$, $O(^1D)$ and O_3. From their associated production-loss equations, the chemical lifetimes of $O(^3P)$, $O(^1D)$ and O_3 are approximately 0.4s, 1s and 2000s respectively at 30km (*Brasseur and Solomon* [1986]). The short lifetimes are a result of fast reactions between the members of the family. If we treat the family as one species, reactions between $O(^3P)$, $O(^1D)$ and O_3 can be treated as an exchange, or partitioning, among family members and not as production and loss terms. This increases the lifetime of the family to the order of weeks as only reactions that cause a net production or loss in O_x are considered in the O_x continuity equation. Three chemical families are considered in CMAT. The first is O_x, consisting of $O(^1D)$ plus $O(^3P)$ (treated as O), and O_3. NO_x is made up from NO and NO_2, and HO_x is made up of H, OH and HO_2. In order to calculate the relative concentrations of each component within a family, partitioning is carried out. This is done by integrating the linearised production-loss differential equations of each family member. The partitioning ratios are determined from the integrated concentrations as will be discussed in more detail below.

Table 2.3 below, taken from *Harris* [2001], lists the species for which CMAT solves.

Transported	Partitioned
Major constituents	
N_2	
O_2	
O_x ___	$O=O(^1D)+ O(^3P)$, O_3
Minor constituents	
$N(^2D)$	
$N(^4S)$	
NO_x ___	NO, NO_2
HO_x ___	H, OH, HO_2
H_2O	
H_2	
CO_2	
CO	
He	
CH_4	
Photochemical equilibrium assumed for $O(^1D)$, H_2O_2 , NO_3 , NO^+, N^+, N_2^+, O_2^+, and O^+	

Table 2.3 Chemical constituents solved for in CMAT

A list of all chemical reactions included in the original CMAT code, along with associated reaction rates and branching ratios can be found in *Harris* [2001]. Updates to the chemical scheme will be discussed in section 3.6.

2.14.2.1 ODD OXYGEN - $O_x = O_3 + O$

The chemical production and loss of the odd oxygen family members in CMAT, as listed in *Harris* [2001], are shown below

$$\frac{dO_3}{dt} = \{(O)^2(O_2)k_{1a} + (O)(O_2)^2 k_{1b} + (O)(O_2)(N_2)k_{1c}\} -$$

$$O_3\{(J_{O3a}+ J_{O3b}) + (H)a_2 + (OH)a_6 + (HO_2)a_{6b} + (O_3)(NO)b_4 + (NO_2)b_9 + (O)k_3\}$$

$$\frac{dO}{dt} = \{(OH)^2 a_{16} + (H)(HO_2)a_{23c} + (N^4S)(O_2)b_7 + (N^2D)(O_2)b_{7b} + (N^2D)(NO)b_{6a} +$$

$$(N^4S)(NO)b_6 + J_{NO2}(NO_2) + J_{NO3a}(NO_3) + (J_{O3a} + J_{O3b})(O_3) + 2(J_{SRB,HZ} +$$

$$J_{SRC})(O_2) + ion\_prod\_O - (O^1D)(H_2)a_{3b} - (O^1D)(H_2O)a_{1b} - (O^1D)(CH_4)c_{1a}\}$$

$$- \mathbf{O}\{(H_2)a_3 + (OH)a_5 + (HO_2)a_7 + (H_2O_2)a_{31} + (CO)(M)a_{37} + (NO_2)b_3 +$$

$$(O)(CH_4)c_3 + (ClO)d_3 + (O_2^2)k_{1b} + (N_2)(O_2)k_{1c} + (O_3)k_3 + ion\_loss\_O\} -$$

$$(\mathbf{O})^2\{(O_2)k_{1a}\}$$

where

$$ion\_prod\_O = (N^+)(O_2)R_4 + (O_2^+)(N^4S)R_8 + (O^+)(O_2)R_{10} + 2(O_2^+)(e^-)R_7 +$$

$$(NO^+)(e^-)R_2 + 0.33(aurqo2)$$

$$ion\_loss\_O = (N^+)R_6 + (N_2^+)R_1 + Ji_O(1 + pe\_O(z)) + aurqo$$

where *aurqo2* and *aurqo* are the ionisation rates due to auroral sources.

pe_O(z) is the ratio of photoelectron ionisation to initial solar photon ionisation at altitude z as described in section 2.14.1.

These equations can be simplified such that

$$\frac{dO_3}{dt} = A - B(O_3) \qquad 2.83$$

$$\frac{dO}{dt} = - X(O)^2 - Y(O) + Z \qquad 2.84$$

where A and B represent production and loss of O_3 respectively, Z is production of O, X and Y are loss of O. Linearising equation 2.83 such that A and B are kept constant, and integrating over a step of h from t to $(t + h)$ seconds gives

$$\int_{O3(t)}^{O3(t+h)} \frac{dO_3}{(A - BO_3)} = \int_{t}^{t+h} dt \qquad 2.85$$

$$\left[\frac{1}{-B}\ln(A - BO_3)\right]_{t}^{t+h} = h \qquad 2.86$$

$$\ln\left[\frac{A - BO_{3(t+h)}}{A - BO_{3(t)}}\right] = -Bh \qquad 2.87$$

$$O_{3(t+h)} = \frac{-(A - BO_{3(t)})e^{-Bh} + A}{B} \qquad 2.88$$

$$O_{3(t+h)} = \frac{A}{B}(1 - e^{-Bh}) + O_{3(t)}e^{-Bh} \qquad 2.89$$

Similarly for equation 2.84

$$\int_{O(t)}^{O(t+h)} \frac{dO}{(-XO^2 - YO + Z)} = \int_{t}^{t+h} dt \qquad 2.90$$

$$\frac{1}{\sqrt{Y^2 + 4XZ}}\left[\ln\left(\frac{-2XO - Y - \sqrt{Y^2 + 4XZ}}{-2XO - Y + \sqrt{Y^2 + 4XZ}}\right)\right]_{t}^{t+h} = h \qquad 2.91$$

let

$$\beta = \left[\frac{-2XO_{(t)} - Y - \sqrt{Y^2 + 4XZ}}{-2XO_{(t)} - Y + \sqrt{Y^2 + 4XZ}}\right] e^{h\sqrt{Y^2+4XZ}} \qquad 2.92$$

therefore

$$O_{(t+h)} = \frac{\beta\left(Y - \sqrt{Y^2 + 4XZ}\right) - \left(Y + \sqrt{Y^2 + 4XZ}\right)}{2X - \beta 2X} \qquad 2.93$$

Equations 2.89 and 2.93 can be used to determine the ratio of O to O_3 and thus the relative concentrations of O and O_3 within the O_x family.

$O(^1D)$ is assumed to be in photochemical equilibrium such that

$$\frac{d(O^1D)}{dt} = 0 = J_{O3a}(O_3) + J_{SRC}(O_2) + J_{H2Ob}(H_2O) - O(^1D)[(H_2O)a_{1b} + (H_2)a_{3b} + (CH_4)c_{1a} + (O_2)k_{4b} + (N_2)k_{4c}]$$

$$O(^1D) = \frac{J_{O3a}(O_3) + J_{SRC}(O_2) + J_{H2Ob}(H_2O)}{(H_2O)a_{1b} + (H_2)a_{3b} + (CH_4)c_{1a} + (O_2)k_{4b} + (N_2)k_{4c}}$$

2.14.2.2 ODD HYDROGEN - HO_X = H + OH + HO_2

The chemical production and loss of the odd hydrogen family members in CMAT, as listed in *Harris* [2001], are shown below

$$\frac{dHO_x}{dt} = 2[\,(H_2O)J_{H2Oa} + (O\,^1D)(H_2)a_{3b} + (O^1D)(H_2O)a_{1b} + (H_2O_2)J_{H2O2}$$

$$+ (O\,^1D)(CH_4)c_{1a} + (O)(H_2)a_3] - 2[\,(HO_2)(H)(a_{23b} + a_{23c}) + (OH)(HO_2)a_{17}$$

$$+ (HO_2)^2 a_{27} + (OH)^2 a_{16} + (OH)(CH_4)c_2]$$

Collecting production and loss terms

$$\frac{dOH}{dt} = - X_{OH}(OH)^2 - Y_{OH}(OH) + Z_{OH}$$

where

$X_{OH} = 2a_{16}$

$Y_{OH} = (O)a_5 + (O_3)a_6 + (HO_2)a_{17} + (H_2)a_{19} + (H_2O_2)a_{30} + (CO)a_{36} + (N^4S)b_{10}$

$Z_{OH} = (H_2O)J_{H2Oa} + (O^1D)(H_2O)a_{1b} + (O_3)(H)a_2 + (O_3)(HO_2)a_{6b} + (HO_2)(O)a_7$

$+ 2(H)(HO_2)a_{23a} + (NO)(HO_2)a_{26} + 2(H_2O_2)\, J_{H2O2} + (O^1D)(H_2)a_{3b}$

$+ (O)(H_2)a_3 + (O^1D)(CH_4)c_{1a} + (H_2O_2)(O)a_{31}$

$$\frac{dHO_2}{dt} = - X_{HO2}(HO_2)^2 - Y_{HO2}(HO_2) + Z_{HO2}$$

where

$X_{HO2} = 2a_{27}$

$Y_{HO2} = (O_3)a_{6b} + (O)a_7 + (H)(a_{23a} + a_{23b} + a_{23c}) + (NO)a_{26} + (OH)a_{17}$

$Z_{HO2} = (H)(O_2)(M)a_1 + (OH)(H_2O_2)a_{30} + (O))(H_2O_2)a_{31} + (OH)(O_3)a_6$

$$\frac{dH}{dt} = - X_H(H)^2 - Y_H(H) + Z_H$$

where

$X_H = 2a_{24}M$

$Y_H = (O_2)(M)a_1 + (O_3)a_2 + (HO_2)(a_{23a} + a_{23b} + a_{23c})$

$$Z_H = (O)(OH)a_5 + (H_2O)J_{H2Oa} + (O^1D)(H_2)a_{3b} + (O)(H_2)a_3 + (OH)(H_2)a_{19} +$$

$$(CO)(OH)a_{36} + (N^4S)(OH)b_{10}$$

The production and loss equations for the members of the HO_x family are partitioned in a similar way to those of the O_x family

$$\frac{OH}{HO_x} = \frac{1}{1 + \frac{H}{OH} + \frac{HO_2}{OH}}$$

Hydrogen peroxide is assumed to be in a state of photochemical equilibrium such that

$$\frac{dH_2O_2}{dt} = 0 = a_{27}(HO_2^{\ 2}) - (H_2O_2)[J_{H2O2} + a_{30}(OH) + a_{31}(O)]$$

therefore

$$H_2O_2 = \frac{a_{27}(HO_2)^2}{[J_{H2O2} + a_{30}(OH) + a_{31}(O)]}$$

2.14.2.3 ODD NITROGEN - NO_x = NO + NO_2

$$\frac{dNO_x}{dt} = [(O_2)(N\ ^2D)b_{7b} + (O_2)(N\ ^4S)b_7 + (NO_3)(J_{NO3a} + J_{NO3b}) + (N^4S)(OH)b_{10}]$$

$$- [\ (NO)J_{NO} + (NO)(N\ ^4S)b_{6b} + (NO)(N\ ^2D)b_{6a} + (NO)Ji_{NO} + (NO_2)(O_3)b_9]$$

$$\frac{dNO}{dt} = A_{NO} - B_{NO}(NO)$$

$$A_{NO} = (NO_2)(O)b_3 + (N^4S)(O_2)b_7 + (N^2D)(O_2)b_{7b} + (N^4S)(OH)b_{10} + (NO_2)J_{NO2}$$

$$B_{NO} = (HO_2)a_{26} + (O_3)b_4 + (N^4S)b_6 + (N^2D)b_{6a} + (O_2^{+})R_9 + J_{NO} + Ji_{NO}$$

$$\frac{dNO_2}{dt} = A_{NO2} - B_{NO2}(NO_2)$$

$A_{NO2} = (HO_2)(NO)a_{26} + (NO)(O_3)b_4 + (NO_3)J_{NO3a}$

$B_{NO2} = (O)b_3 + (O_3)b_9 + J_{NO2} + (NO_3)b_{12}$

Partitioning is carried out as follows

$$\frac{NO_2}{NO_x} = \frac{1}{1 + \frac{NO}{NO_2}}$$

and

$$\frac{NO}{NO_x} = \frac{1}{1 + \frac{NO_2}{NO}}$$

Photochemical equilibrium is assumed for NO_3, therefore in the upper stratosphere

$$\frac{dNO_3}{dt} = 0 = (O_3)(NO_2)b_9 - NO_3[b_{12}(M)(NO_2) + J_{NO3}]$$

$$NO_3 = \frac{(O_3)(NO_2)b_9}{[b_{12}(M)(NO_2) + J_{NO3}]}$$

2.14.2.4 ATOMIC NITROGEN - N(^{4}S)

$$\frac{d(N^4S)}{dt} = ion\_pro_{N4S} + (NO)J_{NO} + (O)(N^2D)b_8 + (N^2D)b_{xx1} + (N^2D)(e^-)R_{13}$$
$$-N^4S\{(O_2)b_7 + (NO)b_6 + (OH)b_{10}\}$$

where ion_pro_{N4S} is the ionospheric production of N^4S

$$ion\_pro_{N4S} = 0.25(Ji_{PN2}) + 0.25(NO^+)(e^-)R_2 + 0.5(2(N_2^+)(e^-)R_3) + (O^+)(N_2)R_{11}$$
$$+ (N^+)(O_2)R_5 + (N^+)(O)R_6 + 0.40((1.34)Pe\_N_2(z)\ Ji_{N2})) + Ji_{N2}(N\_Np(z)$$
$$+ 0.24\ Pe\_N_2(z)\ ) + 0.5(aurqn2*0.24) + 0.25(aurqn2*1.34)$$

Pe_N_2(z) is the ratio of photoelectron ionisation to initial solar ionisation at altitude *z* (see section 2.14.1), *N_Np(z)* the ratio of N^+ production to primary photon ionisation of N_2 (Ji_{N2}) at altitude z, *aurqn2* the ionisation rate due to auroral sources. Ji_{PN2} is the pre-dissociation rate of N_2 and refers to the process where incident photons excite N_2 molecules to *predissociation states* which decay by dissociation of the molecule rather than by radiation. The resulting atomic nitrogen is partitioned between ground state ($N(^4S)$) and excited state ($N(^2D)$) nitrogen. The partitioning factors used in this expression are due to *Fuller-Rowell* [1993].

2.14.2.5 ATOMIC NITROGEN - $N(^2D)$

$$\frac{d(N^2D)}{dt} = ion\_pro_{N2D} - N^2D[(O_2)b_{7b} + (NO)b_{6a} + (O)b_8 + b_{xx1} + (e^-)R_{13}]$$

ion_pro_{N2D} is the ionospheric production of N^2D

$$ion\_pro_{N2D} = (N_2^+)(O)R_1 + 0.75(Ji_{PN2}) + 0.75(NO^+)(e^-)R_2 + 0.5(2(N_2^+)(e^-)R_3)$$
$$+ 0.60((1.34)Pe\_N_2(z)\ Ji_{N2})) + 10Ji_{N2}(N\_Np(z) + 0.24\ Pe\_N_2(z))$$
$$+ 0.5(aurqn2*0.24) + 0.75(aurqn2*1.34)$$

2.14.2.6 WATER VAPOUR – H_2O

$$\frac{d(H_2O)}{dt} = (OH)(OH)a_{16} + (OH)(HO_2)a_{17} + (OH)(H_2O_2)a_{30} + (OH)(H_2)a_{19}$$
$$+ (H)(HO_2)a_{23c} + (CH_4)(OH)c_2 - H_2O[J_{H2Oa} + J_{H2Ob}]$$

2.14.2.7 MOLECULAR HYDROGEN – H_2

$$\frac{d(H_2)}{dt} = (HO_2)(H)a_{23b} + (H_2O)J_{H2Ob} + (H)^2(M)a_{24} - H_2[(OH)a_{19} + (O^1D)a_{3b} + (O)a_3]$$

2.14.2.8 METHANE – CH_4

$$\frac{d(CH_4)}{dt} = -CH_4[(OH)c_2 + (O)c_3 + (O^1D)c_{1a}]$$

2.14.2.9 CARBON DIOXIDE – CO_2

$$\frac{d(CO_2)}{dt} = (CO)(OH)a_{36} + (CO)(O)(M)a_{37}$$

2.14.2.10 CARBON MONOXIDE – CO

$$\frac{d(CO)}{dt} = (CH_4)(O^1D)c_{1a} + (CH_4)(O)c_3 + (CH_4)(OH)c_2 - (CO)[(OH)a_{36} + (O)(M)a_{37}]$$

2.14.2.11 HELIUM – He

$$\frac{d(He)}{dt} = 0$$

2.14.3 ION CHEMISTRY

The ion chemistry code in CMAT is derived from the original CTIM routines of *Fuller-Rowell* [1996a]. The distribution of O^+ is calculated first, then the concentrations of N^+, O_2^+, N_2^+ and NO^+ are determined assuming photochemical equilibrium. $N(^2D)$ and $N(^4S)$ are calculated self-consistently. At latitudes between ±25° the O^+ density and electron and ion temperature are linearly interpolated from the boundary values of the high latitude code. Electron and ion densities are calculated using these interpolated values.

2.14.4 STARTUP CONSTITUENT CLIMATOLOGIES AND BOUNDARY MIXING RATIOS

The start up climatologies used in CMAT are derived from a variety of sources. At altitudes where data is available, O_3, H_2O, N_2O and CH_4 are taken from the UARS reference atmosphere project (*Remedios el al.* [1998]). Above the altitudes covered by UARS, N_2O is linearly extrapolated and global means from the John Hopkins University Applied Physics Laboratory two-dimensional model (JHU/APL 2-D) (*Zhu et al.* [1999]) are used for O_3, H_2O, and CH_4. O, O_2 and H are taken from MSISE-90 and N_2 is calculated as the remainder of the major constituents. Both NO and NO_2 are

made up of monthly zonal mean dawn observations from the UARS HALogen Occultation Experiment (HALOE) instrument. Where HALOE data is not available (above about 56km) NO, is taken as a global mean from the NCAR TIE-GCM (*Roble et al.* [1987]). $N(^4S)$ and $N(^2D)$ are also set as global means from TIE-GCM. CO_2 is assumed to have a constant mixing ratio of 360 ppm up to the turbopause, above which its concentration falls off with its scale height. Global mean profiles are taken from *Roble* [1995] for the remaining constituents.

For the upper two pressure levels, the concentration of H is set to that of MSIS-E90. For all other minor constituents, diffusive equilibrium is assumed at the upper boundary. A full description of the start-up constituent climatologies and boundary mixing ratios used in CMAT can be found in *Harris* [2001].

CHAPTER III. CMAT MODEL MODIFICATIONS AND UPDATES

3.1 INTRODUCTION

This chapter outlines the main updates made to the original CMAT model and their impact on the core model results. Comparisons are made between the original code, the updated code and other numerical models or observations where available.

3.2 PLATFORM COMPATIBILITY

The CMAT model was originally developed on an SGI Origin 2000 comprising of 24 300MHz IP27 MIPS R12000 processors. In order to improve the usability and portability of the code, alterations were made such that CMAT now runs on a cluster of Sun Microsystem SunFire V880 Servers each comprising 8 UltraSparc III 750MHz processors. Interest from the Space Environment Centre in Colorado prompted a further porting of the code onto a Hewlett Packard HP-UX machine. This process was technically non trivial and included the removal of redundant code, multiple bug fixes, the inclusion of error trapping and an all-round improvement to the robustness and usability of the model.

3.3 NEW LOWER BOUNDARY

At the lower boundary of CMAT (set at a level of 10mb pressure, approximately 30km altitude), seasonal geopotential height oscillations are applied in order to correctly represent the middle atmosphere wind structure (see section 2.13.1). In the original version of CMAT, these height oscillations are daily zonal means taken from the Mass Spectrometer and ground based Incoherent Scatter Extended (MSIS-E90) empirical model (*Hedin* [1991]). This empirical model describes the neutral temperature and densities in the Earth's atmosphere from ground to thermospheric heights. The lower boundary temperature is also set from MSIS data, along with several start-up chemical constituent profiles.

In the latest version of CMAT, the MSISE90 model has been replaced with the NRLMSIS-00 empirical model (*Picone et al.* [2002]). This updated model provides a more accurate representation of the neutral atmosphere through the inclusion of

multiple new data sets. Most notably, drag and accelerometer data are now used in the calculation of total mass density which could affect the model coefficients of temperature and composition. New incoherent scatter radar data directly influence the model temperature. UV occultation data from the Solar Maximum Mission is used in an improved calculation of molecular oxygen number density.

3.4 CHAPMAN FUNCTION

Calculations of parameters such as thermospheric heating and photolysis rates require an accurate description of the optical depth of the atmosphere, as described in section 1.3.7. The volume through which solar radiation travels must therefore be known. At large solar zenith angles (>75°), calculation of the atmospheric column of each absorbing species is performed using numerical evaluation of the Chapman Grazing incidence function $Ch(x,\chi)$.

In the original CMAT code, the Chapman function was estimated using approximate expressions due to *Swinder and Gardener* [1967]. The expressions were complex and required cumbersome library routines. These have been replaced by the following simpler, more computationally efficient expressions due to *Smith and Smith* [1972]

$$Ch(x,\chi \le 90°) = \left[\left(\frac{\pi}{2}\right)x\right]^{1/2} e^{y^2} erfc(y) \qquad 3.1$$

$$Ch(x,\chi > 90°) = (2\pi x)^{1/2}\left\{(\sin\chi)^{1/2}\exp[x(1-\sin\chi)] - \frac{1}{2}e^{y^2}erfc(y)\right\} \qquad 3.2$$

where $x=(R+z)/H$, R being the radius of the Earth, z the altitude, and H the scale height of the absorber being considered. χ is the solar zenith angle at the point of interest and y is defined by

$$y = \left(\frac{\pi}{2}x\right)^{1/2} |\cos\chi| \qquad 3.3$$

The error function *erfc(y)* is approximated by

$$erfc(0 \le y \le 8) \approx \frac{1.0606963 + 0.55643831y}{1.0619896 + 1.7245609y + y^2} e^{-y^2} \qquad 3.4$$

$$erfc(8 \le y \le 100) \approx \frac{0.56498823}{0.06651874 + y} e^{-y^2} \qquad 3.5$$

A comparison of the photoionisation rate of O_2 at 250km altitude due to solar photons, as calculated using the *Swinder and Gardener* and newly implemented *Smith and Smith* approximations is shown in Figure 3.1. Conditions are appropriate to northern winter solstice at high solar activity (F10.7 = 190), at a time of 12:00UT. As expected, the peak ionisation occurs over the sunlit southern high latitude region. At lower latitudes, the photoionisation rate drops to zero during the night when no solar photons reach this altitude. The magnitude of the peak ionisation rate is close to $1.8 m^{-3}s^{-1}$ in both plots. The regions of most interest here, are those with high solar zenith angles such as at sunset. While the new *Smith and Smith* approximation results in a reasonably smooth transition from day to night, the *Swinder and Gardener* approximation results in a jagged profile, indicating resolution or implementation problems in the model. Similar profiles are produced by the two approximations for N_2 and O photoionisation rates. The day-night transition is a key period for many chemical elements and as such '*spikey*' profiles of photoionisation are likely to result in inaccuracies in the photochemical scheme. Calculations of electron densities will also be affected. The *Smith and Smith* Chapman function approximation is therefore a significant improvement over that previously used.

3.5 THE SOLAR SPECTRUM AND THERMOSPHERIC HEATING

The original CMAT thermospheric heating scheme, described in section 2.12.3, calculated heating in the UV and EUV regions from 1.8-180nm. 5nm spectral intervals were used for the majority of the wavelength range covered, with some individual line intensities including e.g. Lyman alpha. Tables of fluxes used in the scheme were for two F10.7cm solar flux values, one low and one high, due to *Torr et al.* [1979] and *Torr et al.* [1980c]. These spectra were based on measurements made by spectrometers carried by the Atmospheric Explorer satellites. Fluxes for all other F10.7 values were obtained by linearly interpolating between these two sets of data. Solar radiation at wavelengths below 1.8nm was not considered in the heating calculation, only in the calculation of major species ionisation rates. The omission of hard X-ray flux from the heating calculations was not considered important as the thermal balance of the atmosphere is largely determined by the longer wavelength part of the solar spectral range (*Tobiska et al.* [2000]).

Over recent years it has become evident that EUV models based on some limited set of measurements over a limited period of time can lead to misleading results if applied to a different part of the solar cycle or to a different solar cycle. Inconsistencies exist between commonly used flux models, such as the one used in CMAT, and between models and data. *Lean* [1990] and *Richards et al.* [1994] discuss the problems these types of models have in reproducing both observed photoelectron fluxes and observed EUV flux variations over the solar cycle. *Buonsanto et al.* [1995] compared electron densities calculated with two commonly used EUV models (EUVAV and EUV94X) with measurements made by the incoherent scatter radar at Millstone Hill and found significant differences arose when using different EUV fluxes. The EUV flux is a fundamental parameter for modelling chemistry and dynamics of the thermosphere. Emissions such as Lyman-α (121.6nm), Lyman-β (102.6nm) and hard X-rays, from 0.1 to 0.8nm, dominate energy deposition in the D-region. Therefore an accurate representation of the solar spectrum in CMAT is crucial. A summary of existing solar irradiance measurements and models can be found in *Tobiska et al.* [2000].

It is also important to consider the way the solar spectrum changes with solar cycle. Although the absolute amount of XUV-UV radiation reaching the Earth is much smaller than the amount of longer wavelength radiation, it is the XUV-UV part of the spectrum that varies most over the solar cycle. These emissions originate in the high temperature outer layers of the Sun's atmosphere and are strongly related to magnetic activity. Variability factors over a solar cycle can be 10 to 20 or more between 1 and 10nm, while a factor of 2 is typical for wavelengths between 10 and 120nm. Between 120 and 400nm, solar cycle changes of 1-10% are common. Variability in the TSI (total solar irradiance) is however usually of the order 0.1% (*Tobiska et al.* [2000]). In order to simulate the effects associated with solar cycle variability it is therefore essential to represent the variability of each part of the solar spectrum with as much accuracy as possible. The current method of linear interpolation between two low resolution spectra, one at high and one at low F10.7, is rather crude and in need of updating. In Figure 3.2, the solar flux between 1 and 20nm as given by the SOLAR2000 empirical solar irradiance model (*Tobiska et al.* [2000]) is shown for different values of the F10.7cm solar activity proxy. The photon fluxes used in the old CMAT model, for the same wavelength and activity ranges are also shown. Clearly visible is a large discrepancy between the empirical model and CMAT fluxes, the 1nm resolution of the SOLAR2000 spectrum giving a more detailed representation than the coarse CMAT spectrum. The plot demonstrates the non-linearity of photon flux variation with solar cycle over this highly variable part of the solar spectrum. Between 5.5 and 20nm, the photon flux given by the SOLAR2000 model for an F10.7 of 137 is actually greater than the flux at an F10.7 of 172. In the same wavelength range, the photon flux for an F10.7 of 208 is almost exactly equal to that given for an F10.7 of 243.

In order to increase the resolution, variability and accuracy of the solar spectrum in CMAT, data from the SOLAR2000 model has been included in the CMAT code as detailed below.

3.5.1 THE SOLAR2000 EMPIRICAL IRRADIANCE MODEL

The need to characterise a historical, spectrally self-consistent, empirical solar irradiance spectrum motivated the development of SOLAR2000, an ongoing collaborative project combing multiple sources of data from 1947 to the present day. Versions of the code are currently available in the form of an IDL (interactive data language) program with GUI interface from http://www.spacewx.com/solar2000.html. The model characterises solar irradiance variability across the spectrum from 1-1,000,000nm in 1nm bins as default, although other bin sizes can be specified. Data from satellite missions such as SOLRAD, SNOE, TIMED, GOES, AE-E, SOHO, SME, Nimbus 7 are included along with rocket data and reference spectra as detailed in *Tobiska et al.* [2000].

3.5.2 IMPLEMENTATION OF THE SOLAR2000 MODEL IN CMAT

The data used in CMAT comes from SOLAR2000 Research Grade V1.24a. This version uses empirical data for wavelengths between 1 and 420nm (X-ray, EUV, FUV, UV). At longer wavelengths where photon flux is less variable over solar cycle, the ASTM E-490 solar reference standard is used. Dates covered by the model run from 14th February 1947 to 30th June 2003.

The solar fluxes from SOLAR2000 are tabulated within the CMAT code for typical days covering twelve values of F10.7, ranging from 67 to 243. For wavelengths between 6 and 180nm, bins of 1nm are used. Wavelengths between 1.8 and 6nm are in bins less than 1nm in size. These bins are chosen to correspond to available photoabsorption and ionisation cross section data.

While it is now possible to run CMAT with the spectrum from a specific date or event in the SOLAR2000 database, it is not computationally practical for general use. Typical spectra from 12 solar activity levels, as defined by the F10.7cm flux activity proxy, are therefore used and an interpolation between F10.7 values is used to determine the solar flux at intermediate activity levels.

Figure 3.3 shows a comparison between the new SOLAR2000 spectrum and the old CMAT spectrum for various solar activities, as characterised by F10.7. Clearly seen is an increase in solar flux at X-ray wavelengths in the SOLAR2000 spectrum. The flux of H Lyman-α photons at 121.6nm is larger in the SOLAR2000 data than in the CMAT spectrum, especially at higher values of F10.7. A window in the molecular oxygen cross section allows these photons to penetrate to the mesosphere where they cause the formation of the ionospheric D-region through photoionisation of nitric oxide. An accurate representation of Lyman-α flux and its variability is therefore key to modelling odd nitrogen species (NO_x) and D region chemistry. There is a clear increase in the flux of HeII 30.4nm photons in the SOLAR2000 data over the old CMAT spectrum. Solar radiation at this wavelength constitutes an important source of heating in the thermosphere (*Tobiska et al.* [2000]) and contributes to night-time ionisation of O_2 in the E region (*Banks and Kockarts* [1973]). Increased wavelength resolution means that energy deposition can be more accurately modelled in discrete atmospheric layers.

Bailey [2002] and *Swaminathan* [2001] showed that solar soft X-ray irradiance in the 2-7nm wavelength range has a dramatic effect on the production of nitric oxide, leading to a peak in concentration at around 106km altitude. Radiation in the ranges 7-17nm and 17-20nm was shown to have only a small effect on modelled NO abundances. The solar flux in the 2-20nm range as given by the original CMAT code and by SOLAR2000 is shown in Figure 3.4 for high and low solar activity. Comparison of the high and low activity plots shows that the amount of 2-20nm X-ray flux reaching the Earth is highly variable with solar cycle and increases considerably during solar maximum. From low to high solar activity, the total flux given by the original CMAT spectrum in the 2-20nm range increases by a factor of 3.3, compared to a factor of 4.1 given by SOLAR2000. This suggests that all CMAT elements influenced by the intensity of soft X-ray radiation will display a greater variability with solar cycle using the new SOLAR2000 spectrum.

The SOLAR2000 2-13nm flux is considerably larger than that given by the original CMAT spectrum at both high and low solar activity. Flux in the 13-20nm range is

lower in the SOLAR2000 spectrum than that given by the old CMAT spectrum under both active and quiet conditions. The integrated flux in the 2-7nm range is a factor of 144 (22) larger in the SOLAR2000 spectrum than in the old CMAT spectrum at low (high) solar activity. This is likely to have a dramatic effect on the calculated concentration of NO in the model. The change in flux of radiation in the 7-20nm range is expected to have only a small effect on NO production.

3.5.3 HARD X-RAY FLUXES

Hard X-ray fluxes at wavelengths between 0.2 to 1.8nm are now represented in CMAT and are taken from the GLOW model (*Solomon et al.* [1988], *Solomon and Abreu* [1989]).

3.5.4 UPDATED PHOTOIONISATION AND ABSORPTION CROSS SECTIONS

In order to implement the SOLAR2000 spectrum in CMAT it was necessary to update the photoionisation and absorption cross sections for each of the major species. The new cross sections have been tabulated in bins of 1nm or less. For wavelengths between 0.2 and 6nm, cross sections are based on *Henke* [1993]. 1nm binned cross sections for wavelengths between 6nm and 105nm are based on *Fennelly and Torr* [1992] (*Stan Solomon and Liying Qian* : private communication). Between 105nm and 180nm, cross sections are based on multiple sources (*Rodney Viereck*: private communication).

3.5.5 STEADY STATE CMAT RESULTS

3.5.5.1 TEMPERATURES

Figure 3.5 shows the zonal mean temperatures as given by NRLMSIS-00, CMAT using the old solar spectrum and CMAT using the new high resolution SOLAR2000 spectrum along with updated photoabsorption and ionisation rates. Each plot shows simulated temperature profiles at spring equinox under conditions of low solar and geomagnetic activity. Clearly visible is an increase in CMAT thermospheric temperatures modelled using the new SOLAR2000 spectrum, such that the CMAT temperatures now closely match those given by MSIS. Figure 3.6 shows the zonal

mean temperature plots for similar conditions as Figure 3.5 but at high solar activity. The increase in thermospheric temperatures modelled using the updated solar fluxes and photolysis rates are less pronounced than those observed at low solar activity. The small increase in temperatures does however bring the new CMAT temperatures closer in agreement with the MSIS data over the CMAT profiles derived using the old solar flux data.

Zonal mean plots for temperature calculated at northern winter solstice are given in Figure 3.7 and Figure 3.8 for low and high solar activity, at low geomagnetic activity. As for the equinox plots, the CMAT thermospheric temperatures are in better agreement with MSIS data when the SOLAR2000 solar fluxes and updated photolysis rates are used for both high and low solar activity. The global mean exospheric temperatures (not shown) from MSIS, CMAT with the old solar spectrum and CMAT with the new spectrum are 804K, 632K and 756K respectively at low solar activity. A similar increase is seen in the high solar activity case.

The zonal mean mesopause temperatures given by CMAT at solstice are higher than those given by MSIS by around 20K at both high and low solar activity when SOLAR2000 fluxes are used. This is likely to be caused by the simplistic representation of planetary wave forcing in CMAT which affects the simulated middle atmosphere circulation. MSIS thermospheric temperatures are still slightly higher than those modelled by CMAT using the updated solar fluxes, possibly in relation again to the crude representation of planetary wave drag in CMAT. Other factors could include molecular diffusion coefficients and the high latitude electric field strength. The electric field strength variability applied in CMAT (see section 2.11.4) does not have any dependence on solar cycle, geomagnetic activity or location and as such could contribute to errors in the calculated Joule heating.

3.5.5.2 ELECTRON DENSITIES

Figure 3.9 shows the mid day electron density at a mid latitude location (-54°) in the southern hemisphere as calculated by CMAT using the old solar spectrum, CMAT using the SOLAR2000 spectrum and as given by the International Reference

Ionosphere (IRI) model (*Bilitza* [2001]). Conditions are northern hemisphere spring equinox, low solar activity (F10.7 = 76) and low geomagnetic activity. The plots demonstrate that implementation of the SOLAR2000 spectrum in CMAT results in a large increase in the calculated electron density at low solar activity. The F2 peak given by the IRI model is of the order $6.8x10^{11}m^{-3}$ at an altitude of about 265km. Under similar conditions, the original version of CMAT calculated an F2 peak of $3.2x10^{11}m^{-3}$ at about 210km, compared to a peak of the order $6.7x10^{11}m^{-3}$ at an altitude of 235km, calculated using the updated solar fluxes. E region electron densities modelled by CMAT using both the old and new solar spectrum are up to a factor of 2 greater than those given by IRI but the altitude of the peak is consistent at around 110km.

Figure 3.10 shows similar plots to those in Figure 3.9 but for a mid latitude location in the northern hemisphere (54°) during a period of high solar activity (F10.7 = 180). As for the low activity case, the height of the F2 peak is raised when using the updated solar spectrum, from about 250km to around 280km. The overall magnitude of the CMAT modelled electron densities are not affected by the solar spectrum used. A similar overestimation in the E region densities is seen at high and low solar activity.

Electron density in the atmosphere is highly variable due its dependence on many factors such as solar zenith angle, location, solar and auroral activity etc. It would therefore be unwise to base conclusions about the ability of CMAT to model electron densities on a single set of model and/or observational conditions. The mid latitude location shown has been selected as it reflects typical changes in the CMAT calculated electron density caused by the updated solar fluxes and photoabsorption / ionisation cross-sections. It is therefore reasonable to conclude that under conditions of low solar activity, modelled electron densities are improved by application of the high-resolution SOLAR2000 fluxes and updated photoabsorption and ionisation cross-sections.

The absence of a detailed electron density calculation at low to mid latitudes means that CMAT electron densities are generally in better agreement with observations at mid to high latitudes. Inclusion of a plasmaspheric code such as that included in the

CTIP model (*Millward et al.* [1996], *Fuller-Rowell et al.* [1996]) would improve low latitude electron densities and may improve the overestimated electron densities in the E region.

3.5.5.3 O(^{1}S) GREEN LINE EMISSION

The commonly observed atomic oxygen 557.7 nm green line emission occurs when oxygen in the excited O(^{1}S) state spontaneously relaxes to the lower O(^{1}D) state. The excitation mechanism that precedes this phenomenon will be discussed in detail in Chapter 6, along with the dynamical factors that influence the diurnal variability of the airglow. In CMAT, the green line volume emission rate is calculated using the empirical formulation of *Murtagh et al.* [1990], based on the parameters of *McDade et al.* [1986] as described in section 6.3. This method has been used in numerous modelling studies (e.g. *Yee et al.* [1997], *Roble and Shepherd* [1997], *Zhang et al.* [2001], *Harris et al.* [2002]) in which calculated volume emission ratios were compared to observational data from the Upper Atmosphere Research Satellite (UARS). The magnitude of the simulated airglow in each of these studies was a factor of 2-3 lower than shown in the satellite data. This was attributed to numerous dynamical effects such as an overestimation of the eddy diffusion profile, or underestimation of tidal wind amplitude.

With the updated high-resolution solar spectrum and revised photoabsorption and ionisation rates now included in CMAT, agreement between the magnitude of the simulated and observed green line volume emission rate is improved. It is therefore likely that shortfalls in the magnitude of green line emission rates simulated in the afore mentioned studies are related to the representation of solar flux within the models. Figure 3.11 shows the 557nm volume emission rate at 4LT as a function of latitude and altitude, as calculated by CMAT using the old solar spectrum (top), the updated spectrum (middle) and as given by the WINDII instrument onboard the UARS satellite. The WINDII data is from a composite of longitudinally averaged data that has been binned in intervals of 1 hour and 5 degrees latitude, taken over March and April of 1993 when solar activity was low (*Shepherd et al.* [1995]). The CMAT data is derived from simulations of March equinox under conditions of low solar activity.

Using the old solar spectrum, the maximum in the calculated emission rate in both hemispheres is close to 120 photons $cm^{-3}s^{-1}$, occurring around 98km altitude between 35 and 40° latitude. The structure of the emission profile is largely controlled by the diurnal tide which has been enhanced in the CMAT runs, in line with the study of *Harris et al.* [2002]. This will be discussed in detail in Chapter 6. The peak emission rate calculated by CMAT using the updated solar fluxes and photoabsorption / ionisation rates reaches 180 and 165 photons $cm^{-3}s^{-1}$ in the northern and southern hemispheres respectively. The altitude and latitude at which the peaks occur are approximately the same as in the previous case, although the maxima are more localised when using the new spectrum. Green line emission rates given by WINDII reach a maximum of 170 photons $cm^{-3}s^{-1}$ in the north and 130 photons $cm^{-3}s^{-1}$ in the south. These maxima occur close to ±35° latitude, between 90 and 95km altitude. This is slightly lower in altitude than the maxima given by CMAT. The vertical distribution of the WINDII emission rate is more extended than that simulated by CMAT, with significant amounts of emission occurring below 90km. Differences in the calculated and observed rates are attributed to dynamical influences and the average nature of the WINDII data.

This data present in Figure 3.11 is representative of green line emission at 4LT and was chosen to demonstrate that the magnitude of the green line volume emission rate can vary by more than a factor of two with changing latitude. A strong local time dependence also exists, as demonstrated by *Shepherd et al.* [1995] who showed that at 20LT, the WINDII data reveals a peak emission rate of 200 photons $cm^{-3}s^{-1}$ centred over the equator.

Figure 3.12 shows the 557nm emission rate at the equator as a function of altitude and local time as calculated by CMAT using the two solar spectra, and as given by WINDII. CMAT calculates maxima of 140 and 210 photons $cm^{-3}s^{-1}$ when using the old and new solar spectra respectively. WINDII data shows a maximum of about 200 photons $cm^{-3}s^{-1}$. Note that the CMAT data presented covers a whole day, whereas the WINDII data covers just 6 hours. Both model and data show a minimum in the green line emission rate close to midnight, with enhanced emission in the morning and

evening sectors. This is associated with tidally induced vertical motions of the airglow layer throughout the day, as described by *Shepherd et al.* [1995]. The difference in local time of the calculated and observed oxygen emission peak is thought to be related to a shift in tidal phase, caused by the gravity wave scheme used in CMAT (*McLandress* [1998]). Overall, the magnitudes of the 557nm green line volume emission rate calculated by CMAT using the updated solar inputs are in better agreement with those given by the WINDII satellite than when the original solar inputs were used.

3.6 UPDATED REACTION RATES AND BRANCHING RATIOS

Whilst able to represent the overall morphology of NO_x in the atmosphere, the original version of CMAT was not able to realistically reproduce the observed number densities. *Harris* [2001] noted that the CMAT model was underestimating the nitric oxide concentration by a factor of about 4 to10 in the lower thermosphere, similar to the modelling studies of *Fuller-Rowell* [1993] and *Swaminathan et al.* [1998]. Nitric oxide plays an important role in cooling of the thermosphere and as such it is vital that CMAT be able to model it realistically. In order to improve the representation of NO_x in CMAT, new branching ratios and reaction rates were collated, tested and included in the CMAT chemical scheme. Some additional reactions were added where appropriate. This will be discussed in detail in Chapter 4. The key changes to reaction rates and branching ratios are given in Table 3.1 where chemical equations are listed along with the excess kinetic energy of the reaction. The full CMAT chemical scheme is given in Table 3.2 at the end of this chapter, and APPENDIX A.

	Reaction	Old Coefficient	New Coefficient	Ref.
b_4	$NO + O_3 \rightarrow NO_2 + O_2 + 2.08eV$	$2.0 \times 10^{-18}exp(-1400/T)$	$1.8 \times 10^{-18} exp\ (-1370/T)$	b
b_7	$N(^4S) + O_2 \rightarrow NO + O + 1.4eV$	$4.4 \times 10^{-18}exp(-3220/T)$	$1.5 \times 10^{-17} exp\ (-3600/T)$	l
b_6	$N(^4S) + NO \rightarrow N_2 + O + 2.68eV$	$3.4 \times 10^{-17}exp(120/T)$	$T > 400K$: 3.25×10^{-17} $T < 400K$: $2.2 \times 10^{-17} exp\ (160/T)$	k
b_{7b}	$N(^2D) + O_2 \rightarrow NO + O + 1.84eV$	6.0×10^{-18}	$6.2 \times 10^{-18}(T/300)$	i
b_8	$N(^2D) + O \rightarrow N(^4S) + O + 2.38eV$	4.5×10^{-19}	6.9×10^{-19}	h
b_{8a}	$N(^2D) + O_3 \rightarrow NO + O_2 + 5.43eV$	None	$2.0 \times 10^{-18} T^{1/2} exp(-1200/T)$	j
b_{10}	$N(^4S) + OH \rightarrow NO + H + 2.1eV$	$5.0 \times 10^{-17}exp(-3220/T)$	5.0×10^{-17}	b
b_{12}	$NO_2 + NO_3 \rightarrow N_2O_5 + 0.91eV$	None	$2.2 \times 10^{-36} (T/300)^{3.9}$	a
d_4	$NO + CLO \rightarrow NO_2 + CL$	None	$6.2 \times 10^{-18} exp\ (294/T)$	e
r_3	$N_2^+ + e^- \rightarrow$ $N(^2 D) + N(^2 D) + 5.82eV$ (10%) $N(^4 S) + N(^2 D) + 3.44eV$ (90%)	$1.8 \times 10^{-13}(T_e/300)^{-0.39}$	$1.6 \times 10^{-13}(T_e/300)^{-0.37}$	k
r_4	$N^+ + O_2 \rightarrow NO^+ + O + 6.699eV$	2.0×10^{-16}	2.6×10^{-16}	k
r_{4a}	$N^+ + O_2 \rightarrow O^+ + NO + 2.31eV$	None	3.6×10^{-17}	k
r_5	$N^+ + O_2 \rightarrow O_2^+ + N(^4S) + 2.486eV$	4.0×10^{-16}	3.1×10^{-16}	k
r_{13}	$N(^2 D) + e^- \rightarrow N(^4S) + e^- + 2.38eV$	$3.6 \times 10^{-16}(T_e/300)^{0.5}$	$3.8 \times 10^{-18}(T_e)^{0.81}$	k
J_{N4S}	$N(^4 S) + h\nu \rightarrow N^+ + e^-$	None	$1.0 \times 10^{-7}\ s^{-1}$	g

	Reaction	Old Branching ratio	New Branching ratio	Ref.
r_2	$NO^+ + e^- \rightarrow$			g
	$N(^2D) + O + 0.38eV$	75%	85%	
	$N(^4S) + O + 2.75eV$	25%	15%	
r_3	$N_2^+ + e^- \rightarrow$			k
	$N(^2D) + N(^2D) + 5.82eV$	10%	10%	
	$N(^4S) + N(^2D) + 3.44eV$	90%	90%	
		0.5 $N(^4S)$: 0.5 $N(^2D)$	0.44 $N(^4S)$: 0.56 $N(^2D)$	

Table 3.1 Main changes to reaction rates and branching ratios in CMAT chemical scheme[ID176]. Where not specified, rates are given in $m^{-3}s^{-1}$.

REFERENCE KEY

a - JPL97, *DeMore et al.* [1997]

b - *Roble* [1995]

e - *Brasseur and Solomon* [1986]

g - *Barth* [1992]

h - *Bailey* [2002]

i - *Duff et al.* [2003]

j - *Banks and Kockarts* [1973]

k - *Swaminathan et al.* [1998]

l - *Siskind et al.* [2004]

3.7 NO, NO_2 and $N(^4S)$ START-UP CLIMATOLOGIES

Start-up climatologies of NO_2 and NO, made up of observational data from the UARS reference atmosphere project (see section 2.14.4), are still used in the new version of CMAT at altitudes where data is available. The invariant global mean polynomials, used for NO start up profiles at heights above the lower mesosphere, were found to be highly unrealistic in both magnitude and vertical structure, and have now been replaced. Global data from the Student Nitric Oxide Explorer (SNOE) (*Barth et al.* [2003] and references therein) is now used. This satellite used a limb-scanning technique to take measurements of NO in the mesosphere and lower thermosphere over a two and half year period. The SNOE data provides physical start-up profiles with seasonal and latitudinal variation. The $N(^4S)$ start-up climatology is now taken from MSIS as the original profile was again very crude and invariant with time and space.

By starting the CMAT model with more realistic constituent distributions, the model is able to reach a *stable* self-consistent solution in less time than before, where 'stable' implies that the diurnal variability of a constituent's distribution does not differ significantly from that of the previous day. This reduces the run time of the model. Note that the start-up profiles for NO do not display any diurnal variability or any variation with solar or auroral activity.

3.8 VIBRATIONAL RELAXATION OF NO(v=1) BY O ATOMS

Following excitation by collision with oxygen atoms, 5.3μm infrared emissions from excited nitric oxide causes cooling in the thermosphere. This is a critically important process in determining the thermal budget of the thermosphere. The highly variable concentration of NO leads to a variable cooling rate that probably maximises at high latitudes during auroral NO enhancements (*Kockarts* [1980]).

The process of NO cooling starts when ambient O atoms collide with ground state NO and promote it to its first excited state.

$$NO(\upsilon = 0) + O \rightarrow NO(\upsilon = 1) + O$$

Radiative cooling occurs when a portion of the resultant $NO(\upsilon = 1)$ emits an infrared photon

$$NO(\upsilon = 1) \rightarrow NO(\upsilon = 0) + h\upsilon(5.3\mu m)$$

The deactivation rate coefficient of vibrationally excited NO is therefore an important factor when determining the energy budget of the thermosphere (by convention, the rate coefficients are given in the exothermic direction). Typically a value of 6.5×10^{-11} cm^3s^{-1} (*Kockarts* [1980]) is used to model the photochemistry of the Earth's upper atmosphere after *Fernando and Smith* [1979]. *Dodd et al.* [1999] recently measured the $NO(\upsilon = 1 \rightarrow 0)$ rate experimentally and reported a value of $2.4(\pm 0.5) \times 10^{-11}$ cm^3s^{-1}, close to the theoretically calculated value of 2.8×10^{-11} cm^3s^{-1} given by *Duff et al.* [1997]. *Sharma and Roble* [2002] reviewed the theoretical and experimental status of this rate coefficient and found that varying the rate in a global mean model resulted in dramatic changes in the temperature and density structure of the simulated thermosphere.

Changes made to the chemical scheme and the thermospheric heating scheme in CMAT have lead to a marked increase in the modelled NO concentration, as shall be discussed in Chapter 4. A resultant increase in cooling due to NO radiative emission caused the CMAT thermospheric temperatures to be notably lower than those given by the MSIS model (*Hedin* [1991, *Picone et al.* [2002]). The value of 2.4×10^{-11} cm^3s^{-1} has now been adopted in CMAT, resulting in a good agreement between CMAT and MSIS thermospheric temperatures.

3.9 THE MEDVEDEV AND KLASSEN GRAVITY WAVE SCHEME

As previously mentioned, global scale models of the atmosphere use parameterisations of the effects of gravity waves on the large-scale atmospheric circulation because the spatial resolution required to model individual waves is prohibitively fine. The original version of CMAT uses a *Lindzen-Matsuno* scheme as described in section 2.13.3. This scheme is able to reasonably recreate the observed zonal mean wind structure in the middle atmosphere under certain conditions. It is however rather simplistic, ignoring

the effects of wave-wave interactions. *Hamilton* [1997] found that Lindzen type schemes, which treat multiple waves independently, tend to produce drag in GCMs at altitudes that are too high if moderate launch amplitudes are used. Using stronger launch amplitudes induces momentum deposition at lower altitudes but in this case the drag is too strong. The scheme used in CMAT is finely tuned to simulate momentum deposition of an acceptable strength at a reasonable height and uses an efficiency parameter which acts to reduce the drag, based on the assumption that wave forcing is intermittent. *Harris et al.* [2002] suggested that use of this parameterisation in CMAT resulted in a reduction in the amplitude of the diurnal tide, as shall be discussed in Chapter 6. The simplicity of this scheme may also contribute to the poor representation of the winter pole through an incomplete treatment of winter mesospheric gravity wave forcing. Perceived shortfalls in the parameterisation have lead to the implementation of a new gravity wave scheme based on *Medvedev and Klaassen* [1995, 2000, 2003] (hereafter MK2000). The *Medvedev and Klaassen* scheme was provided by Scott England [*private communication* 2004]. The choice of this parameterisation shall be discussed further in Chapter 6.

The MK2000 gravity wave parameterisation employs a modification of *Weinstock's* [1990, 1993] theory of nonlinear (i.e. spectrum-induced) wave diffusion whereby the low-frequency part of the gravity wave spectrum is treated as an additional background flow for high-frequency waves.

The scheme considers the amplitude damping that 'background' spectral components with small vertical scale (high vertical wavenumber m) and longer periods (low frequency ω) exert on a given wave component. A wave of similar or larger scale vertical variations propagating through this background will experience nonlinear interactions with the spectrum. The result is an increase in damping as the wave is forced closer to an overturned state. Damping associated with these interactions can result in saturation or even destruction of the given wave.

The scheme employs the dispersion relation

$$m = \frac{N}{(c - \bar{u})} \tag{3.6}$$

where m is the vertical wavenumber, c is the horizontal phase speed, N is the buoyancy frequency and $\bar{u}(z)$ is the mean wind.

Medvedev and Klaassen [1995] showed that the evolution of gravity wave spectra with height z is given by,

$$\frac{dS(m)}{dz} = \left(-\frac{\rho_{o_z}}{\rho_o} + \frac{m_z}{m} - \beta \right) S(m) \tag{3.7}$$

where $S(m,z)$ is the power-spectral density (PSD) of the horizontal wind associated with gravity waves at altitude z. The PSD is defined as the squared modulus of the Fourier Transform of the time series, or the amount of power per unit frequency as a function of frequency. ρ_0 is the mean density and β is the coefficient of nonlinear damping due to interactions of the component m with other waves in the spectrum. This can be approximated by

$$\beta = \sqrt{2\pi} N \sigma^{-1} \exp(-\alpha^2) \tag{3.8}$$

where N is the buoyancy frequency and the dimensionless parameter α is defined by

$$\alpha = \frac{N}{\sqrt{2} m \sigma} = \frac{c - \bar{u}}{\sqrt{2}\sigma} \tag{3.9}$$

where N is the buoyancy frequency, and $\sigma(m)$ is the horizontal RMS wind created by all waves in the spectrum with vertical wavenumbers larger than the given m. $\sigma(m)$ is determined by

$$\sigma^2 = \int_m^\infty S(m')dm' \tag{3.10}$$

where σ^2 is the horizontal wind variance.

Equations 3.7 to 3.10 describe the vertical evolution of the wave spectra in terms of the vertical wavenumber m alone.

The horizontal drag due to a given spectral component ($a_h(m)$) can now be found from

$$a_h(m) = -\frac{\beta(m)k_h S(m)}{m} \qquad 3.11$$

where $k_h = |\boldsymbol{k}_h|$ is a characteristic horizontal wave number. The total drag is calculated by integrating the contributions from all parts of the spectrum. Thus vertical profiles of gravity wave drag can be determined.

A full description of the numerical implementation of this gravity wave parameterisation can be found in *Medvedev and Klaassen* [2000].

3.9.1 THE GRAVITY WAVE SOURCE SPECTRUM

In order to make the MK2000 gravity wave parameterisation suitable for use in atmospheric models, the continuous input source power spectral density, S, is replaced by a discrete spectrum of waves. This isotropic spectrum consists of 60 waves with 15 frequencies each, in the northern, southern, eastern and western directions. A model spectrum which approximates the typical amplitude and shape of observed horizontal wind vertical wavenumber spectra (*Medvedev* [2000], *Medvedev et al.*, [1997], *Fritts and VanZandt*, [1993]), is obtained by using a modified Desaubies spectrum of the form

$$S(m) = S_0 \frac{(m/m_*)^s}{1+(m/m_*)^{s+t}} \qquad 3.12$$

where $S(m)$ is the power spectral density of each wave, S_0 is the amplitude of the spectrum and m is the vertical wavenumber of each wave. m_* is the '*characteristic wavenumber*' which corresponds to the 'knee' of the observed spectrum. Waves with vertical wavenumbers less than m_* exhibit exponential amplitude growth with height and are said to be unsaturated. Waves with high vertical wavenumbers exhibit an m^{-3}

tail whose intensity is largely unvarying with height. This suggests that waves with high m are being dissipated and as such these waves are described as saturated (*McLandress* [1998]). The parameter t is equal to 3 and defines the 'm^{-3}' power law of the high wavenumber tail of observed saturated spectra (e.g. *Fritts and VanZandt* [1993]).

After *Medvedev and Klaassen* [2000] we assume the modified Desaubies form with s = 1, m_* = 0.006 m^{-1} (e.g. *Fritts and Lu*, [1993]) and values for m range from $2\pi/900$ m^{-1} to $2\pi/19$ km^{-1}. A characteristic horizontal wavenumber (k_h) of $-2\pi/300$ km^{-1} is used. Corresponding phase velocities range from approximately 2 to 60ms^{-1}. A latitudinally varying source amplitude of the form $S_0 = 50(1+100\sin^2 2\theta)$ is used, where θ is latitude, such that the source amplitude maximises at mid-latitudes (*Medvedev et al.* [1997]).

Figure 3.13 shows the power spectral density as a function of horizontal phase speed for the source spectrum at the CMAT lower boundary (10mb). The *Medvedev et al.* [1997] spectrum at 100mb (about 15km) is also shown. The CMAT spectrum was obtained by propagating the *Medvedev et al.* [1997] spectrum from 100mb through a zonally averaged mechanistic model of the atmosphere derived from the Extended Stratosphere Model (*Arnold and Robinson* [1998], [2000], *Scott England* [Private communication 2004]). A value of $2\pi/8$ minutes has been assumed for the Brunt-Väisälä frequency and the background wind speed is set to 0 ms^{-1}. The amplitude of the spectrum launched at 100mb increases with altitude as a result of the exponential decrease in atmospheric density. Both spectra have the same overall shape with the familiar m^{-3} tail at high wavenumbers. Differences between the two spectra arise from dissipation and spectral changes caused by wave-wave interactions between 10mb (15 km) and 100mb (30 km).

Note that zero phase speed waves, sometimes called orographic waves, are not included in the gravity wave spectrum. These waves are usually produced by stationary geological features such as mountains and tend to have large amplitudes, causing them to dissipate in the lower stratosphere. For this reason they are not

expected to have a large impact in CMAT (*Scott England private communication* [2005]).

3.9.2 SELF CONSISTENT EDDY DIFFUSION PROFILE

An expression that enables the self-consistent calculation of vertical eddy diffusion profile was described by *Medvedev and Klaassen* [2003]. This expression is now included in the CMAT scheme and takes the following form

$$D_z = \sum_{j=1}^{M} \frac{a_{hj}}{2m_{Rj}N} \qquad 3.13$$

D_z is the vertical eddy diffusion coefficient in m^2s^{-1}, N is the buoyancy frequency and m_{Rj} is the real part of the vertical wavenumber. j=1 and j=M are the components with the lowest and highest vertical wavenumbers in the gravity wave spectrum respectively and a_{hj} is the horizontal wave drag associated with the jth component, as described by equation 3.11.

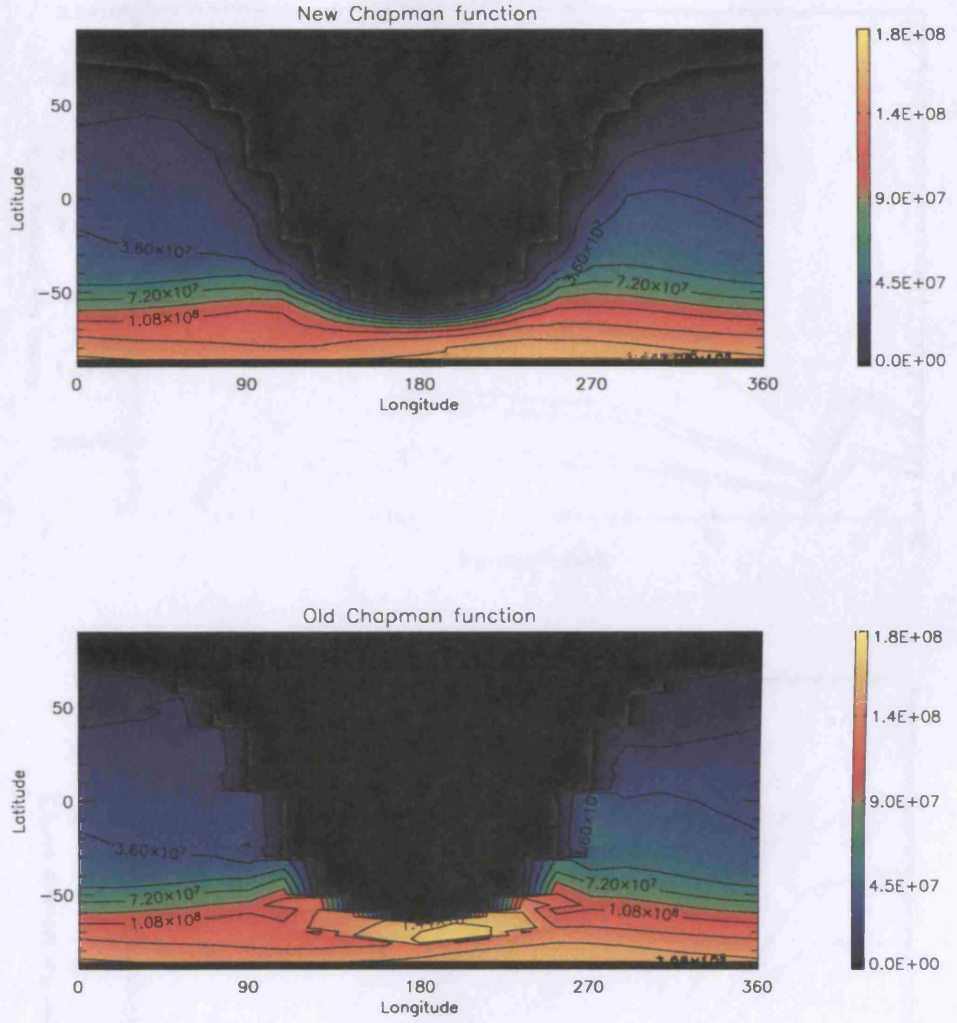


Figure 3.1 O_2 photoionisation rate ($m^{-3}s^{-1}$) at 250km altitude as calculated by CMAT using the *Smith and Smith* [1972] (*top*) and *Swinder and Gardener* [1969] (*bottom*) approximations of the Chapman function. Conditions are northern winter solstice, high solar activity (F10.7 = 190).

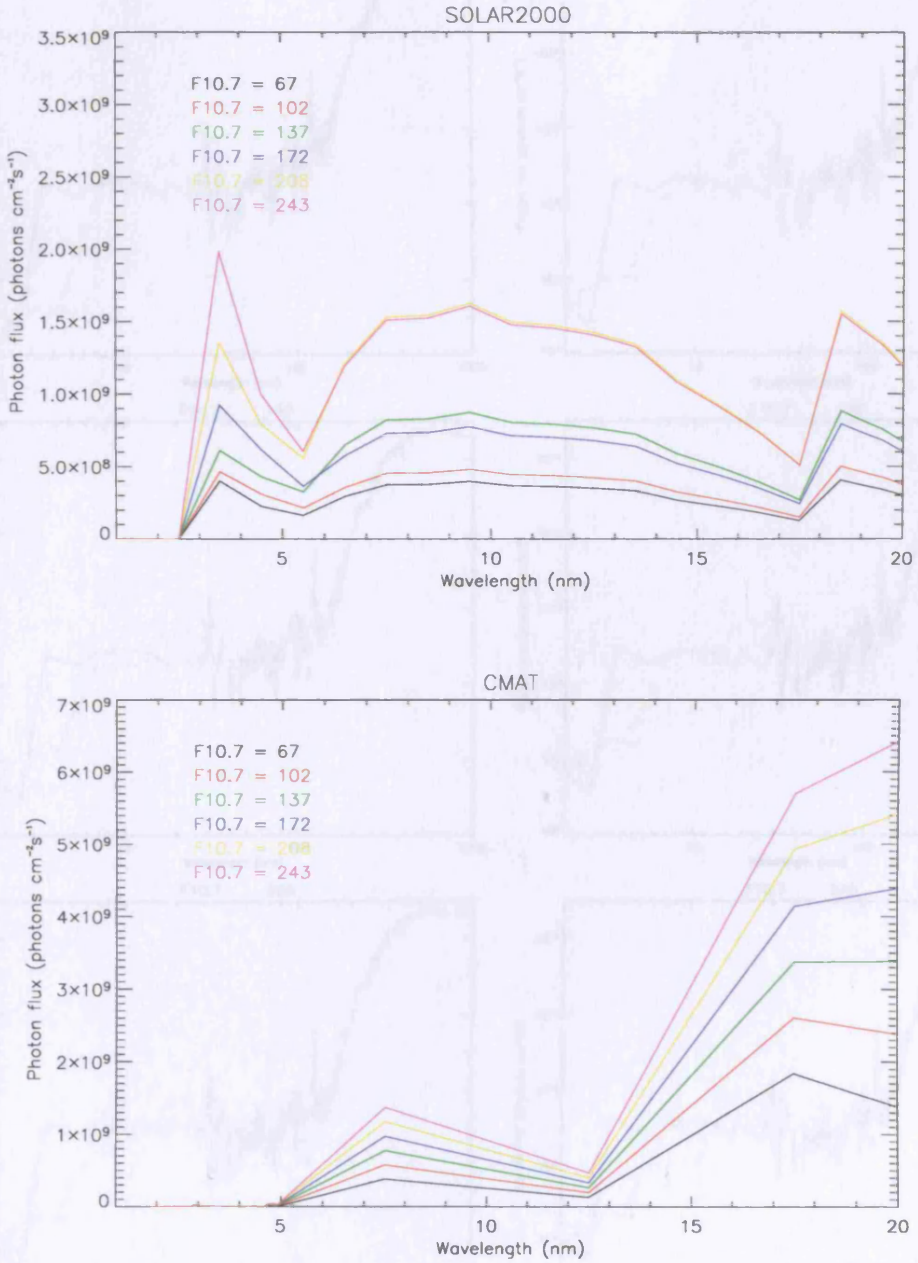


Figure 3.2 Solar photon flux in the 1-20nm range as given by the SOLAR2000 (*top*) and CMAT (*bottom*) models for various levels of solar activity. Note the change of scale between the plots.

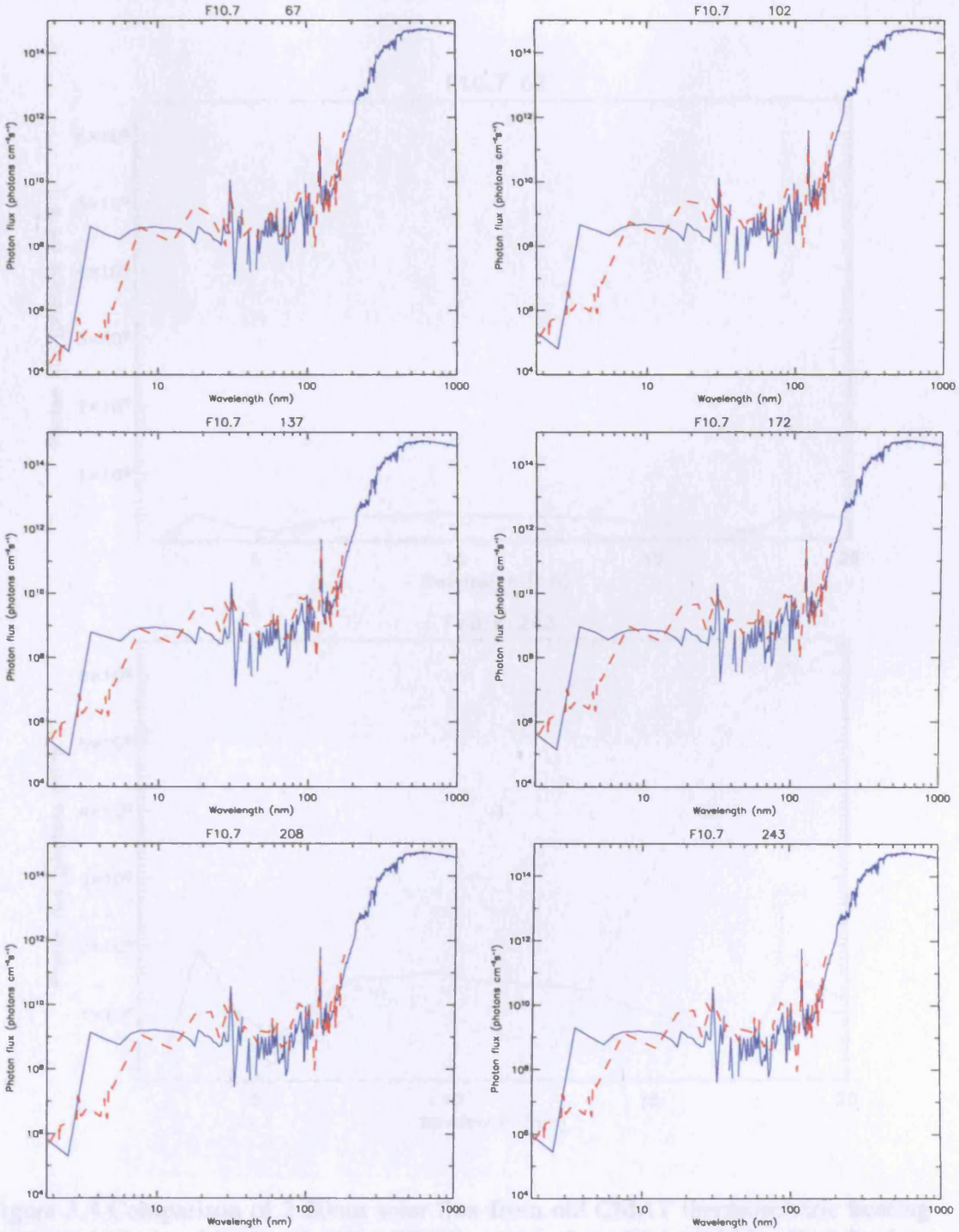


Figure 3.3 Comparison of solar flux from old CMAT thermospheric heating routine (*dashed red line*) and SOLAR2000 empirical model in 1nm bins (*solid blue line*) for different values of F10.7.

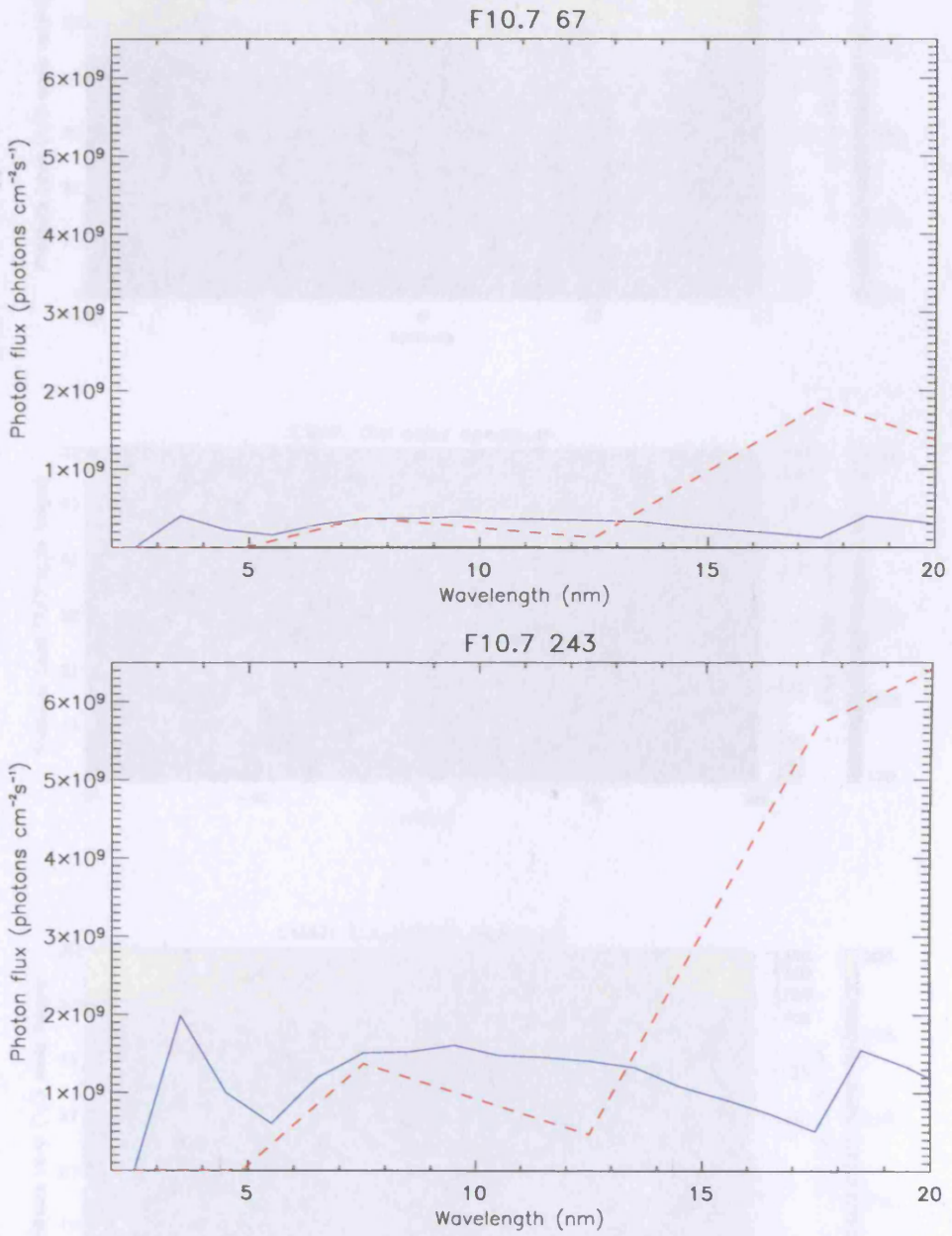


Figure 3.4 Comparison of 2-20nm solar flux from old CMAT thermospheric heating routine (*dashed red line*) and SOLAR2000 empirical model (*solid blue line*) for low solar activity (F10.7=67) and high solar activity (F10.7=243).

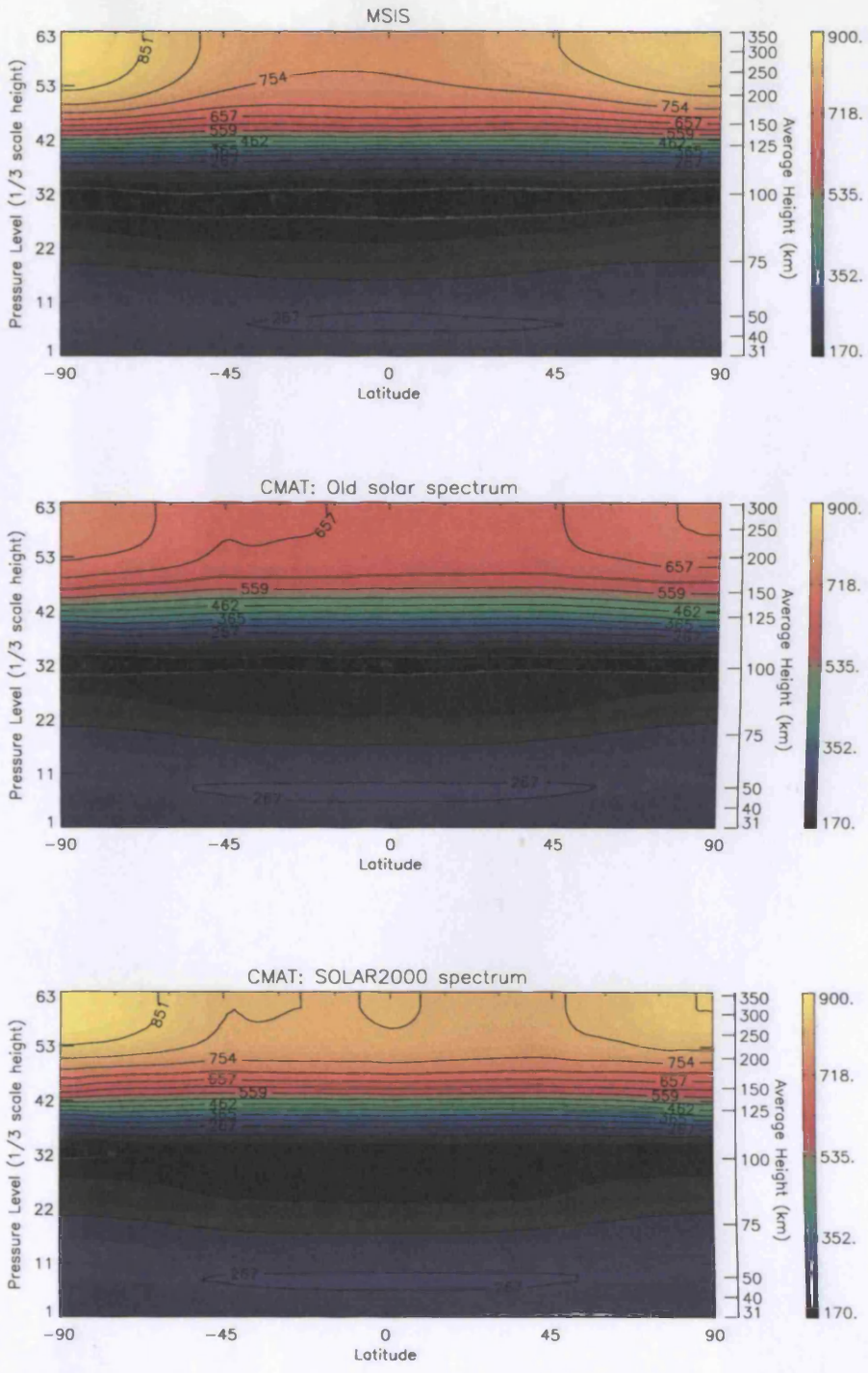


Figure 3.5 Zonal mean Temperature (K) at spring equinox from MSIS (*top*), CMAT using the old solar spectrum (*middle*) and CMAT using the SOLAR2000 solar spectrum (*bottom*). F10.7 =76, K_p=2$^+$.

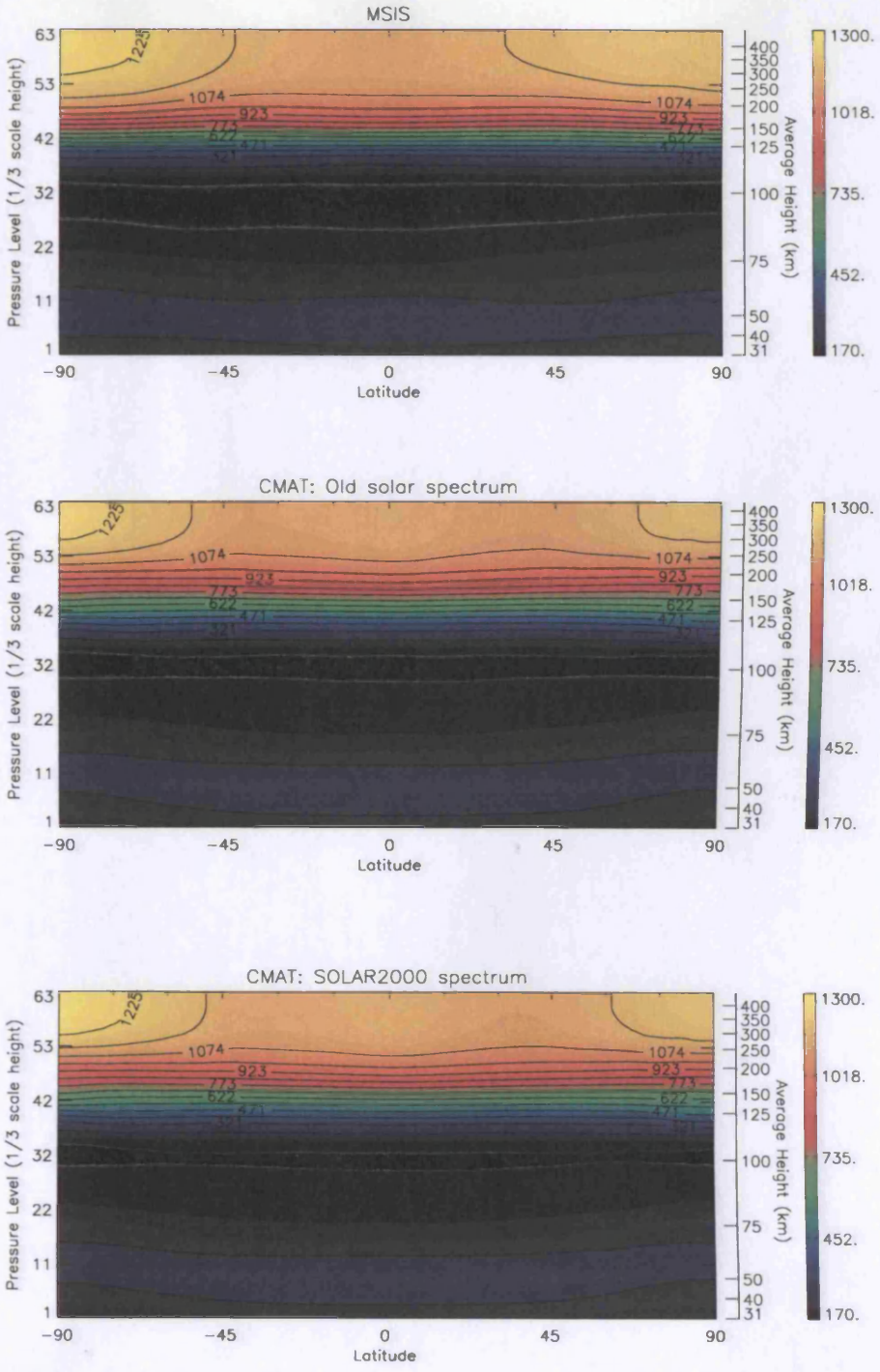


Figure 3.6 Zonal mean Temperature (K) at spring equinox from MSIS (*top*), CMAT using the old solar spectrum (*middle*) and CMAT using the SOLAR2000 solar spectrum (*bottom*). F10.7 =180, K_p=2$^+$.

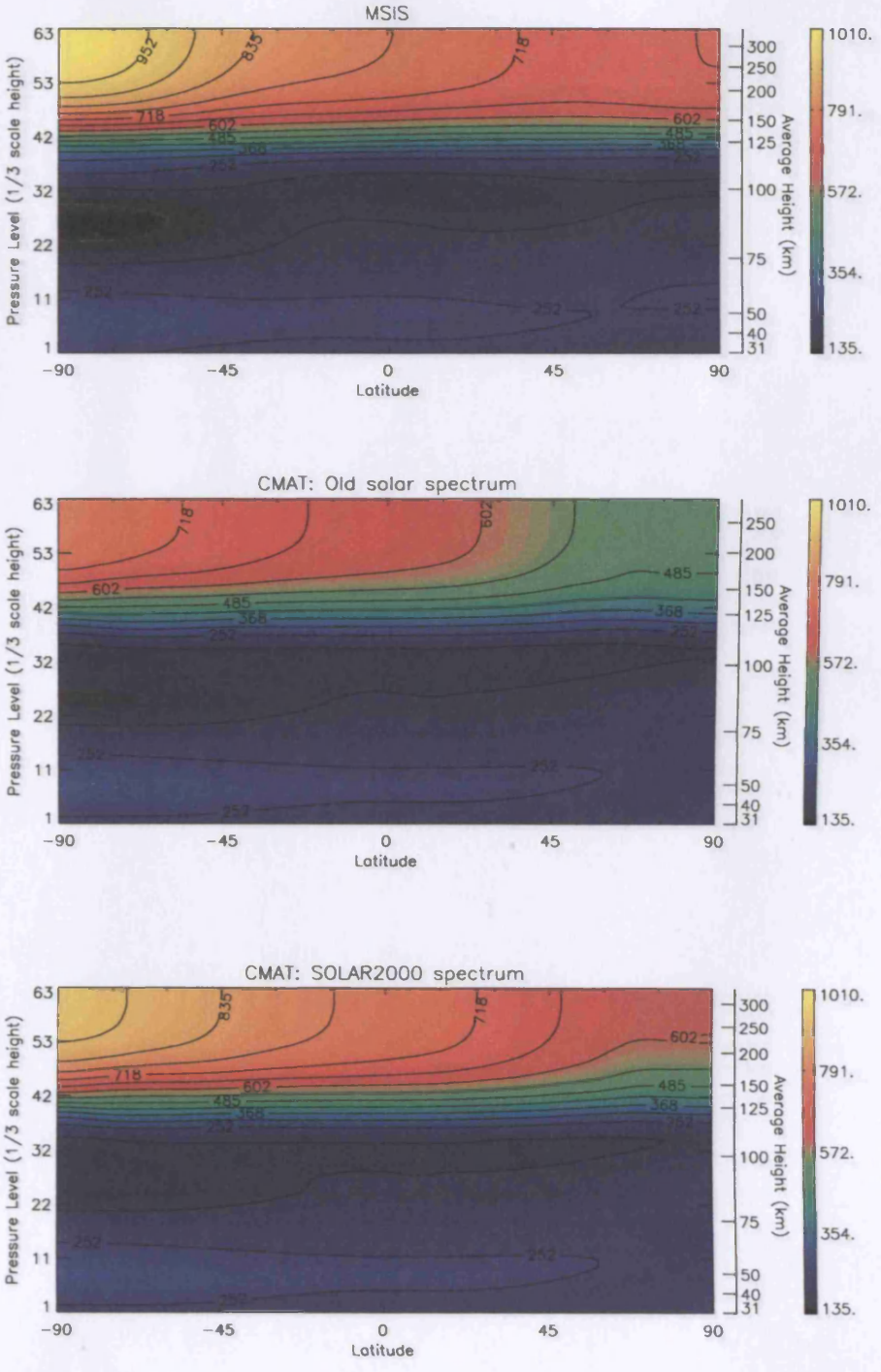


Figure 3.7 Zonal mean Temperature (K) at northern winter solstice from MSIS (*top*), CMAT using the old solar spectrum (*middle*) and CMAT using the SOLAR2000 solar spectrum (*bottom*). F10.7 =76, K_p=2$^+$.

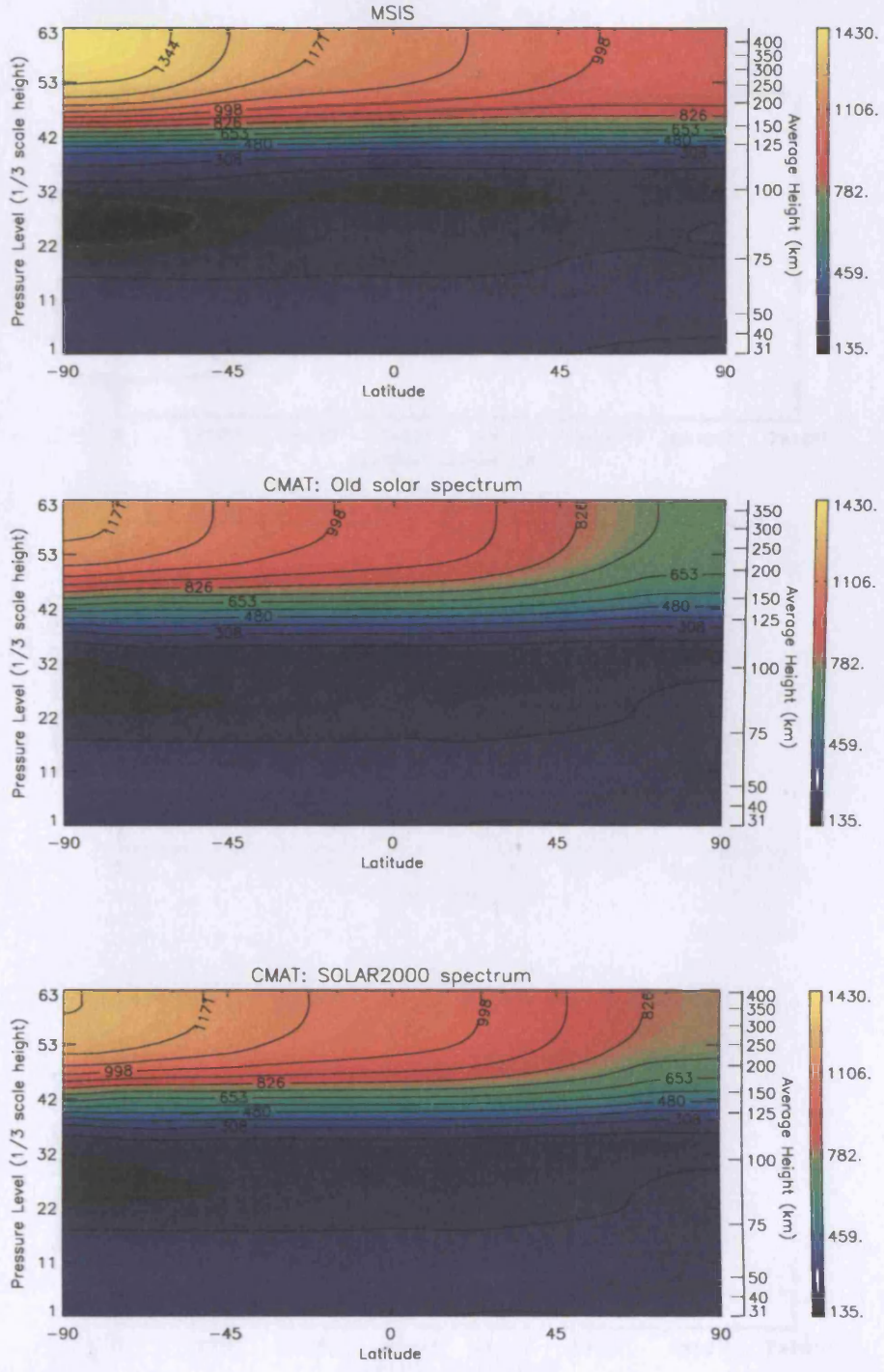


Figure 3.8 Zonal mean Temperature (K) at northern winter solstice from MSIS (*top*), CMAT using the old solar spectrum (*middle*) and CMAT using the SOLAR2000 solar spectrum (*bottom*). F10.7 =180, K_p=2$^+$.

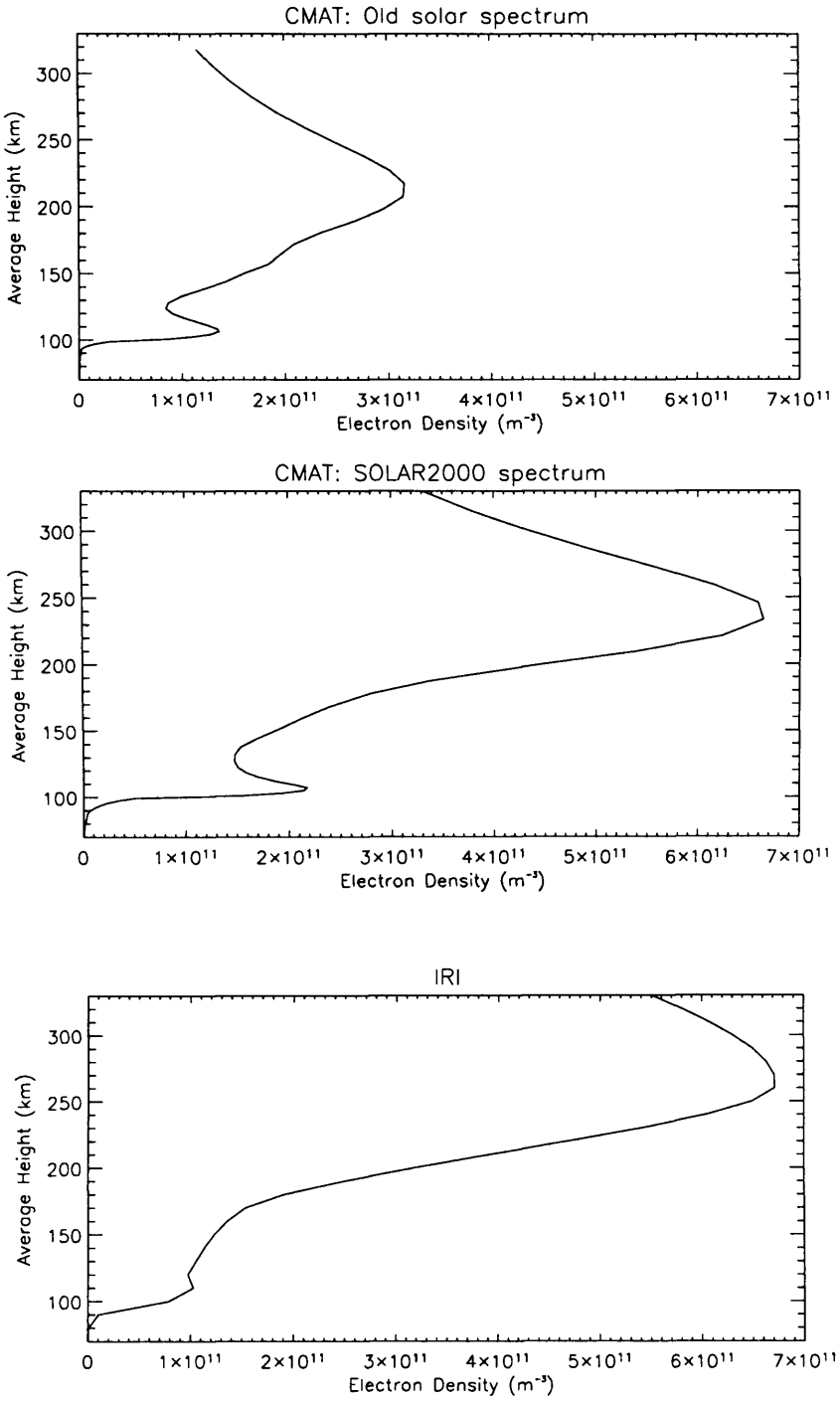


Figure 3.9 Midday, mid latitude electron density vs. height at northern spring equinox as given by the original version of CMAT (*top*), CMAT with the SOLAR2000 solar spectrum (*middle*) and IRI (*bottom*). F10.7 = 76.

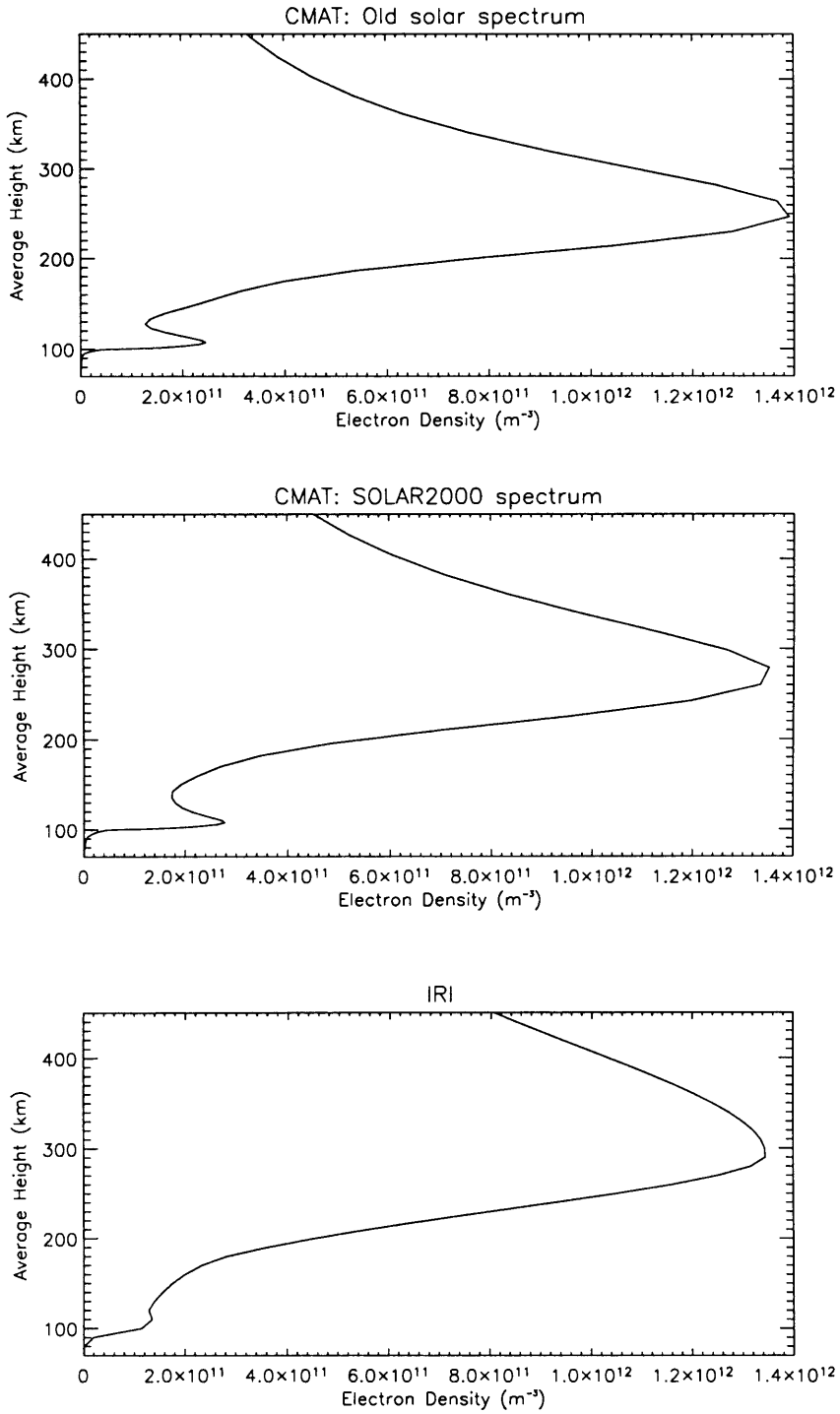


Figure 3.10 Midday, mid latitude electron density vs. height at northern spring equinox as given by the original version of CMAT (*top*), CMAT with the SOLAR2000 solar spectrum (*middle*) and IRI (*bottom*). F10.7 = 180.

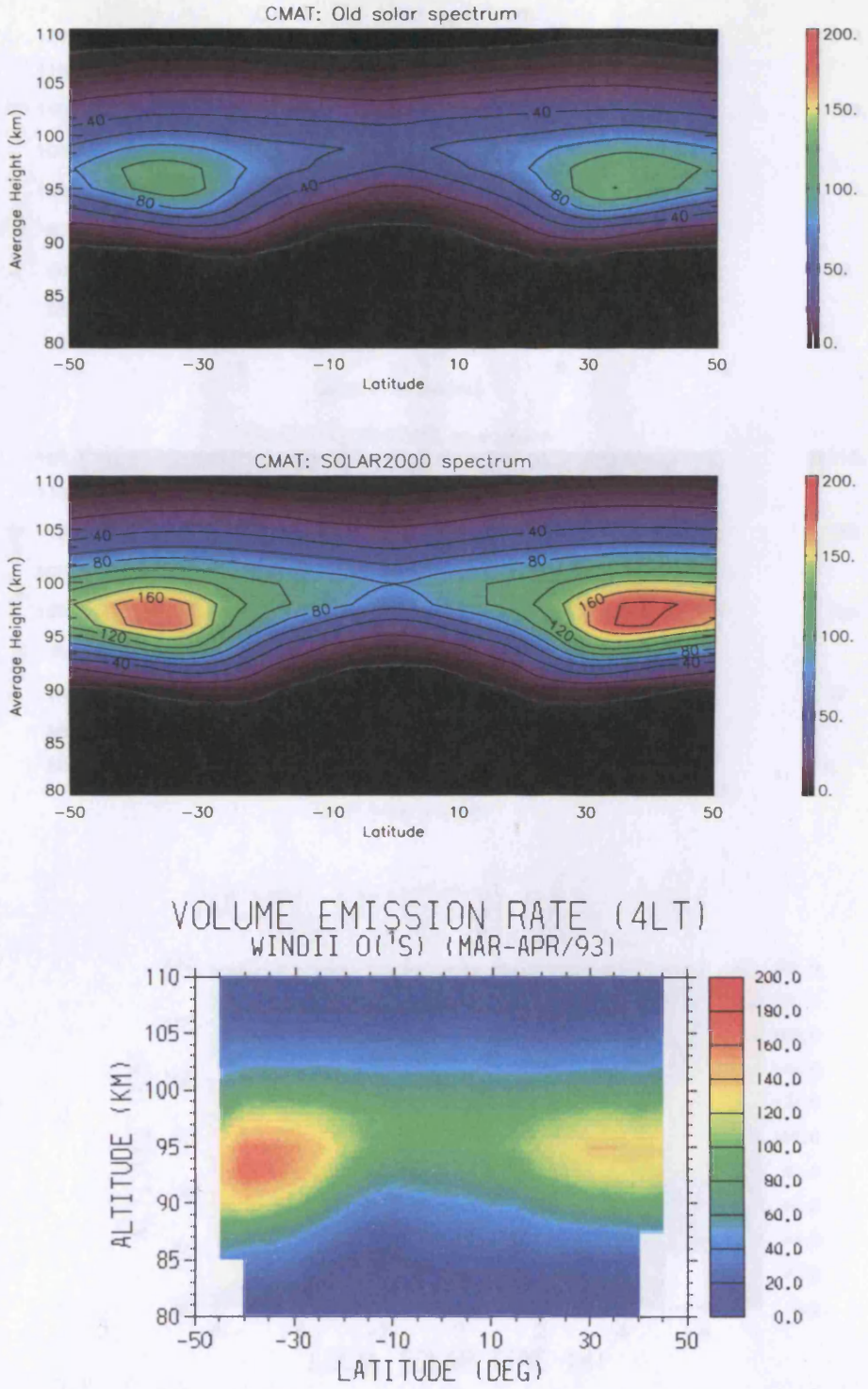
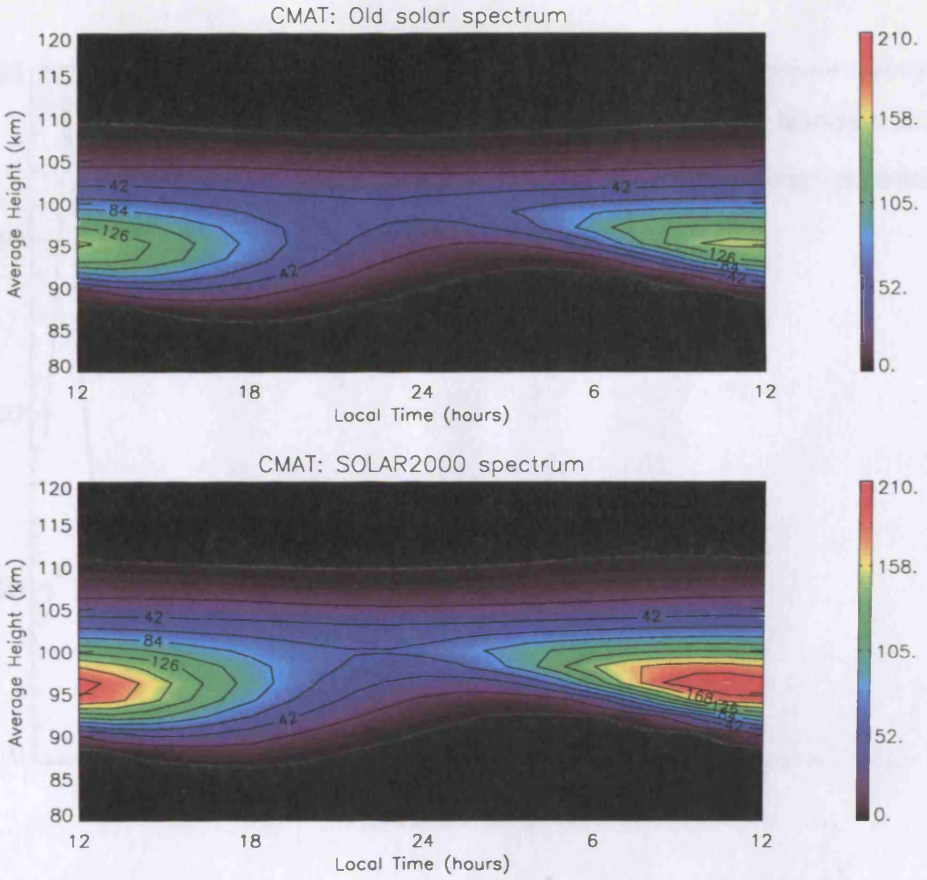


Figure 3.11 Atomic oxygen 557nm volume emission rate (photons $cm^{-3}s^{-1}$) at 4LT as a function of latitude vs. altitude, as calculated by CMAT using the old solar spectrum (*top*), CMAT using the SOLAR2000 spectrum (*middle*) and as given by WINDII (*bottom*). Conditions appropriate to March equinox.



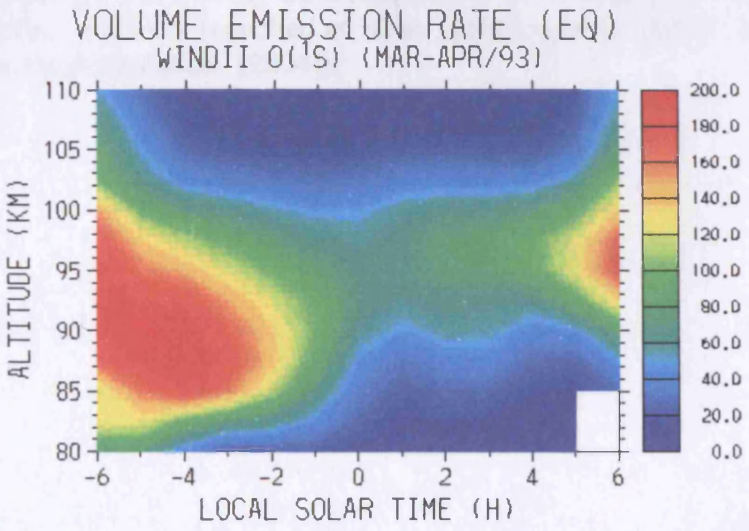


Figure 3.12 Altitude vs. local time plot of atomic oxygen 557nm volume emission rate (photons $cm^{-3}s^{-1}$) as calculated by CMAT using the old solar spectrum (*top*), the SOLAR2000 spectrum (*middle*) and as given by WINDII (*bottom*). Conditions appropriate to March equinox.

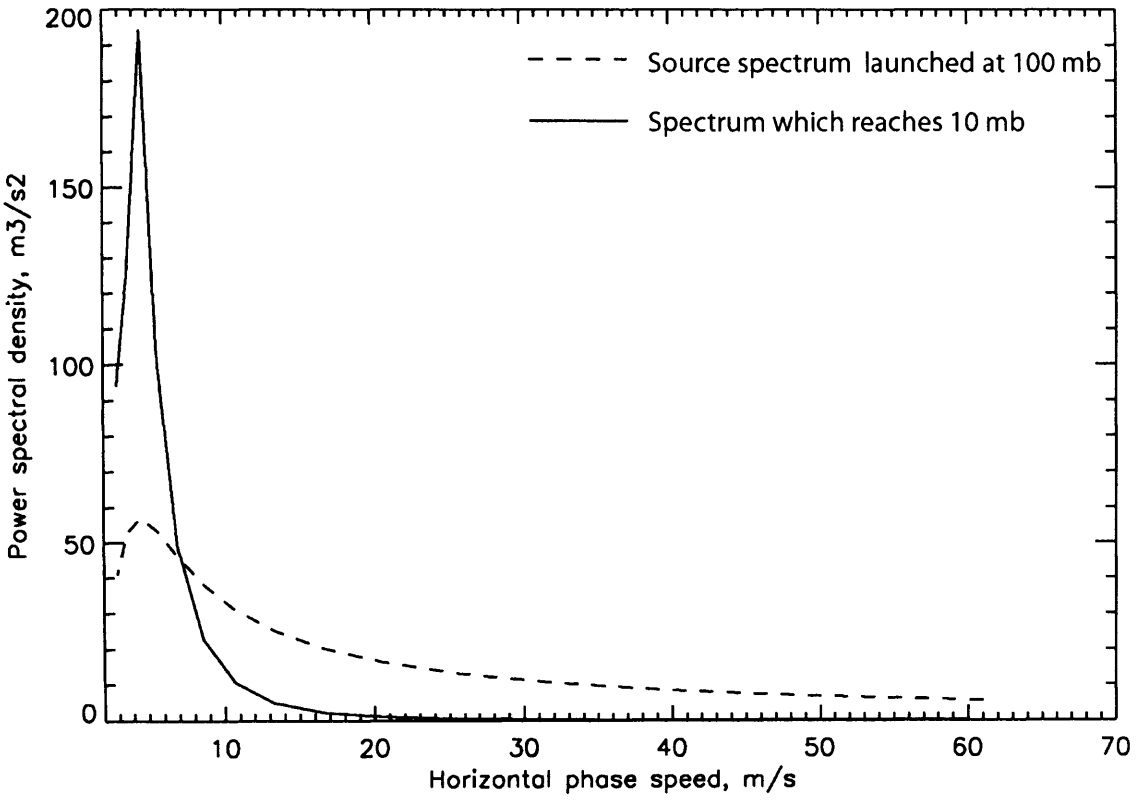


Figure 3.13 Power spectral density as a function of horizontal phase speed for the gravity wave source spectrum launched at 10mb (15km) and at 100mb (30km), (*Scott England*, private communication, [2004]).

Table 3.2 CMAT Chemical scheme

	Reaction			Coefficient	Ref.
T, Te and Ti represent the neutral, electron and ion gas temperatures respectively.					
a_1	$H + O_2 + M$	→	$HO_2 + M + 2.11eV$	$5.7 \times 10^{-44}(T/300)^{-1.6}$	a
a_{1b}	$O(^1D) + H_2O$	→	$OH + OH + 1.23eV$	2.2×10^{-16}	a
a_2	$H + O_3$	→	$OH (v=1\text{-}9) + O_2 + 1.23eV$	$1.4 \times 10^{-16}\exp(-470/T)$	a
a_3	$H_2 + O$	→	$H + OH + 0.08eV$	$1.6 \times 10^{-17}\exp(-4570/T)$	a
a_{3b}	$H_2 + O(^1D)$	→	$H + OH + 1.88eV$	1.0×10^{-16}	b
a_5	$OH + O$	→	$O_2 + H + 0.72eV$	$2.2 \times 10^{-17}\exp(120/T)$	a
a_6	$OH + O_3$	→	$O_2 + HO_2 + 1.73eV$	$1.6 \times 10^{-18}\exp(-940/T)$	a
a_{6b}	$HO_2 + O_3$	→	$O_2 + O_2 + OH + 1.23eV$	$1.1 \times 10^{-18}\exp(-500/T)$	a
a_7	$HO_2 + O$	→	$OH + O_2 + 2.33eV$	$3.0 \times 10^{-17}\exp(200/T)$	a
a_{16}	$OH + OH$	→	$H_2O + O + 0.73eV$	$4.2 \times 10^{-19}\exp(-240/T)$	a
a_{17}	$OH + HO_2$	→	$H_2O + O_2 + 3.06eV$	$4.8 \times 10^{-17}\exp(250/T)$	a
a_{19}	$OH + H_2$	→	$H_2O + H + 0.65eV$	$5.5 \times 10^{-18}\exp(-2000/T)$	a
a_{23b}	$H + HO_2$	→	$H_2 + O_2 + 2.41eV$	$4.2 \times 10^{-17}\exp(-350/T)$	b
a_{23a}	$H + HO_2$	→	$OH + OH + 1.61eV$	$4.2 \times 10^{-18}\exp(-950/T)$	b
a_{23c}	$H + HO_2$	→	$H_2O + O + 2.34eV$	$8.3 \times 10^{-17}\exp(-500/T)$	b
a_{24}	$H + H + M$	→	$H_2 + M + 4.52eV$	$5.7 \times 10^{-44}(T/300)^{1.6}$	b
a_{26}	$HO_2 + NO$	→	$NO_2 + OH + 0.35eV$	$3.5 \times 10^{-18}\exp(250/T)$	a
a_{27}	$HO_2 + HO_2$	→	$H_2O_2 + O_2 + 1.71eV$	$2.3 \times 10^{-19}\exp(600/T)$	a
a_{30}	$OH + H_2O_2$	→	$HO_2 + H_2O + 1.35eV$	$2.9 \times 10^{-18}\exp(-160/T)$	a

a_{31}	$O + H_2O_2$	→	$OH + HO_2 + 3.44eV$		$1.4 \times 10^{-18}exp(-2000/T)$	a
a_{3b}	$O(^1D) + H_2$	→	$OH + H + 1.88eV$		1.1×10^{-16}	a
a_{36}	$CO + OH$	→	$CO_2 + H + 1.07eV$		$1.47 \times 10^{-13}(1 + 0.59P_{atm})$	b
a_{37}	$CO + O + M$	→	$CO_2 + M + 5.351eV$		$6.6 \times 10^{-33}exp(-1103/T)$	b
b_3	$NO_2 + O$	→	$NO + O_2 + 1.98eV$		$6.5 \times 10^{-18}exp(120/T)$	a
b_4	$NO + O_3$	→	$NO_2 + O_2 + 2.08eV$		$1.8 \times 10^{-18}exp(-1370/T)$	b
b_6	$N(^4S) + NO$	→	$N_2 + O + 2.68eV$	$T > 400°K$	3.25×10^{-17}	k
				$T < 400°K$	$2.2 \times 10^{-17}exp(160/T)$	k
b_{6a}	$N(^2D) + NO$	→	$N_2 + O + 5.63eV$		6.7×10^{-17}	c
b_7	$N(^4S) + O_2$	→	$NO + O + 1.4eV$		$1.5 \times 10^{-17}exp(-3600/T)$	l
b_{7b}	$N(^2D) + O_2$	→	$NO + O + 1.84eV$		$6.2x10^{-18}(T/300)$	i
b_8	$N(^2D) + O$	→	$N(^4S) + O + 2.38eV$		6.9×10^{-19}	h
b_{8a}	$N(^2D) + O_3$	→	$NO + O_2 + 5.43eV$		$2. \times 10^{-18} T^{1/2} exp(-1200/T)$	j
b_9	$NO_2 + O_3$	→	$NO_3 + O_2 + 1.08eV$		$1.2 \times 10^{-19}exp(-2450/T)$	a
b_{xx1}	$N(^2D)$	→	$N(^4S) + h\nu$		1.06×10^{-5}	c
b_{10}	$N(^4S) + OH$	→	$NO + H + 2.1eV$		5.0×10^{-17}	b
b_{12}	$NO_2 + NO_3$	→	$N_2O_5 + 0.91eV$		$2.2 \times 10^{-36} (T/300)^{3.9}$	a
b_{13}	$N_2O + O(^1D)$	→	$2(NO) + 2.04eV$		6.7×10^{-17}	b
d_4	$NO + CLO$	→	$NO_2 + CL$		$6.2 \times 10^{-18}exp(294/T)$	e
c_{1a}	$CH_4 + O(^1D)$	→	$CO + 2OH + H_2O + 6.61eV$		1.4×10^{-16}	b

c_2	$CH_4 + OH$	→	$CO + OH + 2H_2O + 5.36eV$	$2.45 \times 10^{-18}\exp(-1775/T)$	a
c_3	$CH_4 + O$	→	$CO + 2OH + H_2O + 6.61eV$	$3.5 \times 10^{-17}\exp(-4550/T)$	b
k_1	$O + O + M$	→	$O_2 + M + 5.12eV$	$9.59 \times 10^{-46}\exp(480/T)$	b
k_{1a}	$O + O_2 + O$	→	$O_3 + O + 1.10eV$	$6.00 \times 10^{-46}(T/300)^{-2.8}$	b
k_{1b}	$O + O_2 + O_2$	→	$O_3 + O_2 + 1.10eV$	$6.00 \times 10^{-46}(T/300)^{-2.8}$	b
k_{1c}	$O + O_2 + N_2$	→	$O_3 + N_2 + 1.10eV$	$6.00 \times 10^{-46}(T/300)^{-2.8}$	b
k_2	$O + O_2 + M$	→	$O_3 + M + 1.10eV$	$6.00 \times 10^{-46}(300/T)^{-2.3}$	a
k_3	$O + O_3$	→	$2O_2 + 4.06eV$	$8.00 \times 10^{-18}\exp(-2060/T)$	a
k_{4a}	$O(^1D) + N_2$	→	$O + N_2 + 1.96eV$	$1.8 \times 10^{-17}\exp(107/T)$	b
k_{4b}	$O(^1D) + O_2$	→	$O + O_2 + 1.96eV$	$3.2 \times 10^{-17}\exp(67/T)$	b
R_1	$O + N_2^+$	→	$NO^+ + N(^2D) + 0.70eV$ $T<1500°K$	$1.4 \times 10^{-16}(300/T_i)^{0.44}$	c
			$T>1500°K$	$5.2 \times 10^{-17}(T_i/300)^{0.2}$	
R_2	$NO^+ + e^-$	→	$N(^2D) + O + 0.38eV$ (85%)	$4.2 \times 10^{-13}(300/T_e)^{0.85}$	h
			$N(^4S) + O + 2.75eV$ (15%)		
R_3	$N_2^+ + e^-$	→	$N(^2D) + N(^2D) + 5.82eV$ (10%)	$1.6 \times 10^{-13}(T_e/300)^{-0.37}$	k
			$N(^4S)_{(0.44)} + N(^2D)_{(0.56)} + 3.44eV$ (90%)		
R_4	$N^+ + O_2$	→	$NO^+ + O + 6.699eV$	2.6×10^{-16}	k
R_{4A}	$N^+ + O_2$	→	$O^+ + NO + 2.31eV$	3.6×10^{-17}	k
R_5	$N^+ + O_2$	→	$O_2^+ + N(^4S) + 2.486eV$ (100%)	3.1×10^{-16}	k
			$O_2^+ + N(^2D) + 0.1eV$ (0%)		c
R_6	$N^+ + O$	→	$O^+ + N(^4S) + 0.98eV$	1.0×10^{-18}	c

R_7	$O_2^+ + e^-$	→	$O + O + 4.98eV$	$T<1200^oK$	$1.6 \times 10^{-13}(300/T_e)^{0.55}$	c
				$T>1200^oK$	$2.7 \times 10^{-13}(300/T_e)^{0.7}$	
R_8	$O_2^+ + N(^4S)$	→	$NO^+ + O + 4.21eV$		1.2×10^{-16}	c
R_9	$O_2^+ + NO$	→	$NO^+ + O_2 + 2.813eV$		4.4×10^{-16}	c
R_{10}	$O_2 + O^+$	→	$O_2^+ + O + 4.865eV$		2.82×10^{-17}– $7.74\times10^{-18}(T_1/300)$ + $1.073 \times 10^{-18}(T_1/300)^2$ – $5.17 \times 10^{-20}(T_1/300)^3$ + $9.65 \times 10^{-22}(T_1/300)^4$	c
R_{11}	$N_2 + O^+$	→	$NO^+ + N(^4S) + 1.088eV$		(i) 1.533×10^{-18} – $5.92 \times 10^{-19}(T_2/300)$ $+ 8.6 \times 10^{-20}(T_2/300)^2$	c
			(i) $300^oK<T_2<1700^oK$			
			(ii) $1700^oK<T_2<6000^oK$		(ii) 2.73×10^{-18} – $1.155 \times 10^{-18}(T_2/300)$ $+ 1.483 \times 10^{-19}(T_2/300)^2$	c
R_{12}	$N_2^+ + O_2$	→	$N_2 + O_2^+ + 3.52eV$		$5.1 \times 10^{-17}(T/300)^{-0.8}$	c
R_{13}	$N(^2D) + e^-$	→	$N(^4S) + e^- + 2.38eV$		$3.8 \times 10^{-18}(T_e)^{0.81}$	k

See reference key for algorithms

$J_{SRB,HZ}$	$O_2 + h\nu$ (λ>242nm)	→	$O + O$	*S.R.B., Hz.*	d(i)
J_{SRC}	$O_2 + h\nu$ (λ<175nm)	→	$O + O(^1D)$	*S.R.C. (inc. Ly α)*	flux calc
J_{O3a}	$O_3 + h\nu$ (λ>310nm)	→	$O_2 + O$	*Hugg., Chapp.*	e(i)
J_{O3b}	$O_3 + h\nu$ (λ<310nm)	→	$O_2 + O(^1D)$	*Hart.*	flux calc
J_{H20a}	$H_2O + h\nu$ ($\lambda\approx$121nm)	→	$H + OH$	*Ly α*	e(ii)

J_{H20b}	$H_2O + h\nu$ (λ>150nm)	→	$H_2 + O(^1D)$	*S.R.B.*	e(iii)
J_{NO2}	$NO_2 + h\nu$ (190-410nm)	→	$NO + O$	1.0×10^{-2}	b
J_{NO3a}	$NO_3 + h\nu$ (λ<580nm)	→	$NO_2 + O$	0.18	f
J_{NO3b}	$NO_3 + h\nu$ (λ>580nm)	→	$NO + O_2$	0.022	f
J_{NO}	$NO + h\nu$ (λ>150nm)	→	$N(^4S) + O$	S.R.B.	d(ii)
J_{N4S}	$N(^4S) + h\nu$	→	$N^+ + e^-$	$1.0 \times 10^{-7}\ s^{-1}$	g
Ji_{NO}	$NO + h\nu$ ($\lambda \approx 121$nm)	→	$NO^+ + e^-$	*Ly* α	e(iv)
$aurq_O$	$O +$ *part.precip.* e_a^*	→	$O^+ + 2e^-$		c
$aurq_{N2}$	$N_2 +$ *part.precip.* e_a^*	→	$N_2^+ + 2e^*$	$(76\%)Q_{AURN2}$	c , k
			$N^+ + N(^4S) + 2e^*$	$(12\%)Q_{AURN2}$	
			$N^+ + N(^2D) + 2e^*$	$(12\%)Q_{AURN2}$	
			$N(^2D)_{0.54} + N(^4S)_{0.46} + (e^*)$	$(134\%)Q_{AURN2}$	
$aurq_{O2}$	$O_2 +$ *part.precip.* e_a^*	→	$O_2^+ + 2e^*$	$(67\%)\ Q_{AURO2}$	c
			$O^+ + O + 2e^*$	$(33\%)\ Q_{AURO2}$	
Ji_{O2}	$O_2 + h\nu$ (λ<105nm)	→	$O_2^+ + e^*$	$(67\%)\ Q_{O2}$	flux calc
	$+ e_p^*$		$O^+ + O + e^*$	$(33\%)\ Q_{O2}$	
Ji_O	$O + h\nu$ (λ<105nm)	→	$O^+ + e^*$		flux calc
	$+ e_p^*$		$O^+ + 2e^*$		
Ji_{N2}	$N_2 + h\nu$ (λ<105nm)	→	$N_2^+ + e^*$ (*or* $2e^*$)	$(76\%)Q_{N2}$	flux calc
	$+ e_p^*$		$N^+ + N(^4S) + e^*$ (*or* $2e^*$)	$(12\%)Q_{N2}$	
			$N^+ + N(^2D) + e^*$ (*or* $2e^*$)	$(12\%)Q_{N2}$	
			$N(^2D)_{0.5} + N(^4S)_{0.5} + (e^*)$	$(134\%)Q_{N2}$	

TABLE REFERENCE KEY

Chemical equations are listed with the excess kinetic energy of each reaction and the rate coefficient in $m^{-3}s^{-1}$. J_X and Ji_X represent the rate of dissociation or ionisation by solar radiation respectively. *Lyα* is the flux of solar Lyman α radiation at zero optical depth taken from the SOLAR2000 reference spectrum ($cm^{-2}s^{-1}$). *S.R.B.*, *S.R.C.*, *Hz*, *Hugg* and *Chapp* represent solar radiation in the Schumann-Runge Bands, the Schumann-Runge Continuum, the Hertzberg continuum, the Huggins bands and the Hartley bands respectively. All daytime photolysis and dissociation rates are functions of solar zenith angle either through the integrated column density along the solar ray path, or are set to constant values for all solar zenith angles then to zero after sunset for rates corresponding to optically thin media (e.g.J_{NO2}). Rates are also a function of solar activity (as characterised by F10.7) where applicable. Night-time ionisation rates are defined below. Two body reactions are in units of m^3s^{-1}, three body m^6s^{-1}, and photo-reactions s^{-1}.

$aurq_X$ represents ionisation of species *X* by auroral particle precipitation where e^* denotes an energetic electron, e_a^* an auroral electron and Q_{AURN2} the ionisation rate due to auroral particles. e_p^* denotes a photoelectron and N_X is the column density of species *X* (cm^{-2}).

This table has been adapted and updated from *Harris* [2001].

a - JPL97, *DeMore et al.* [1997]

b - *Roble* [1995]

c - *Fuller-Rowell* [1993]

d - (i) *Nicolet* [1984]

$$J_{SRB} = [1 + 0.11(F10.7-71)/(220-71)]x$$
$$1.1 \times 10^{-7}\exp[-1.97\times10^{-10}N_{O2}^{0.522}]\ s^{-1}$$

$J_{HZ} = 8.0 \times 10^{-10}\ s^{-1}$ for $z > 45$km, and

$8.0 \times 10^{-10} - [(45 - z)/15](7 \times 10^{-10})\ s^{-1}$ for $z < 45$ km

(ii) *Nicolet* [1979]

$J_{NO} = 4.5 \times 10^{-6}\exp[-10^{-8}(N_{O2})^{0.38}]\exp[-5 \times 10^{-19}N_{O3}]\ s^{-1}$

Night-time $J_{NO} = 1 \times 10^{-6}$[daytime J_{NO} when $\sec\chi=1$]

e - *Brasseur and Solomon* [1986],

$J_{O3a} = [1.2 \times 10^{-4}(Huggins) + 4.4 \times 10^{-4}(Chappuis)]\ s^{-1}$

$J_{H2Oa} = [4 + 2.5(F10.7-71)/(220-71)]\exp[-4.4 \times 10^{-19}N_{H2O}]\ s^{-1}$

$J_{H2Ob} = 1.2 \times 10^{-6}\exp[-1 \times 10^{-7}N_{H2O}^{0.35}]\ s^{-1}$

$Ji_{NO} = (Ly\alpha)\exp[-1x10^{-20}N_{O2}]\ s^{-1}$

Night-time $Ji_{NO} = 1 \times 10^{-6}$[daytime Ji_{NO} at $\sec\chi=1$]

f - *Magnota and Johnston* [1980]

g - *Barth* [1992]

$J_{N4S} = 1x10^{-7}\ s^{-1}$

Night-time $J_{N4S}=0$

h - *Bailey* [2002]

i - *Duff et al.* [2003]

j - *Banks and Kockarts* [1973]

k - *Swaminathan et al.* [1998]

l - *Siskind et al.* [2004]

flux calc Calculated using solar fluxes and absorption cross-section as discussed in section 1.3.8. For O, O_2, and N_2, night-time ionisation = 1×10^{-6}[daytime rate at $\sec\chi=1$]. For ionisation by photoelectrons, the electron profiles are taken from *Fuller-Rowell* [1993] as described in section 2.14.1.

CHAPTER IV. MODELLING OF NITRIC OXIDE IN THE LOWER THERMOSPHERE

4.1 INTRODUCTION

In recent years, increasing attention has been paid to the abundance and variability of Nitric Oxide (NO) in the lower thermosphere. Due to its radiative and chemical properties, this minor constituent plays an important role in the structure and energetics of the upper atmosphere. One reason for interest in this species arises from its low ionisation threshold (ionisation potential 9.26eV). NO concentration has a direct impact on the ion composition of both the E region and the D region, the latter being largely created by photoionisation of NO by solar Lyman alpha (*Nicolet and Aikin* [1960]). Secondly, radiative emission at 5.3μm from NO is an important cooling mechanism in the thermosphere (*Kockarts* [1980]), capable of reducing the heating rate efficiency near 130km by a factor of 4 in the summer hemisphere (*Richards et al.* [1982]). During the polar night, the lifetime of NO is sufficiently long that it can be transported downwards into the stratosphere (*Barth et al.* [1999b], *Callis et al.* [1998], *Vitt et al.* [2000]) where it increases the abundance of odd nitrogen species (NO_y) and can contribute to the catalytic destruction of ozone (*Callis et al.* [2001]). This could affect the temperature structure of the stratosphere and is thought to be one mechanism by which coupling of the upper and lower atmosphere occurs. Finally, NO in the MLT region displays a large degree of variability with season, solar cycle, solar rotation, and auroral activity. Both observational and numerical modelling studies have found a correlation between the flux of solar soft X-rays in the 2-10nm region and low latitude NO concentrations (*Barth et al.* [1988], *Siskind et al.* [1990], *Barth et al.* [1999b], *Swaminathan et al.* [2001]). X-rays in this wavelength band are absorbed between 100 and 120km and the resultant ionisation creates photoelectrons which contribute substantially to NO production. At high latitudes, auroral energy is the primary driver for producing NO (*Barth et al.* [2003], *Baker et al.* [2001], *Solomon et al.* [1999]). As such, a relationship exists between peak NO abundances and auroral activity indicators, such as the K_p planetary geomagnetic index (*Marsh et al.* [2004]). Through consideration of these mechanisms, NO abundance can be used as a diagnostic of the

temporal and spatial distribution of solar and auroral energy deposition into the MLT region.

Despite considerable effort over the last decade, some fundamental questions about the creation, destruction and transport of NO remain unanswered. The profile of NO at any given time evolves through a complex set of ion-neutral and neutral-neutral reactions, many of which are temperature dependent (e.g. $N(^4S) + O_2 \rightarrow NO + O$, *Fuller-Rowell* [1993], *Siskind and Rusch* [1992], *Siskind et al.* [2004]). *Siskind and Rusch* [1992] found that changing the temperature dependence of the reaction between ground state atomic nitrogen and molecular oxygen, in line with JPL90 recommendations, led to an over estimation of NO at 200km by up to a factor of 3 compared to rocket measurements. Ambiguity of key reaction rates and branching ratios combined with sensitivity to highly variable auroral and solar energy inputs means that the modelling of NO production is a nontrivial task. Calculating the loss of NO can be problematic due to uncertainties about the opacity of NO to solar radiation, which can act as a primary destruction mechanism through the dissociation reaction, $NO + h\nu \rightarrow N + O$. Indirect effects such as Joule heating at auroral latitudes can cause enhancements in NO through the temperature sensitivity of the NO production reactions (*Siskind et al.* [1989a,b]). Conflicting measurements of solar soft X-ray have lead to uncertainties in low latitude production rates. Several sensitivity studies have been carried out which investigate how uncertainties in the NO chemical scheme and the state of the natural atmosphere affect the production of NO (*Siskind et al.* [1989a,b], *Barth* [1992], *Siskind et al.* [1995]). Few modelling studies however, have been able to accurately reproduce the absolute magnitudes seen in satellite or rocket data.

Modelling of NO is notoriously difficult and despite recognised demand (e.g. *Barth and Bailey* [2004], *Siskind* [2000]) a comprehensive 3-D model of NO does not, as yet, exist. Without 3D GCM type studies, questions about the spatial and temporal extent of NO transport, both horizontally and vertically, could remain unanswered. Questions about NO creation in key areas where measurement is difficult, such as the winter pole, could also be addressed with 3D GCM studies, as could the timescales of thermospheric cooling following periods of high auroral activity.

In this chapter, recent advances in both observation and modelling of NO are reviewed. A one-dimensional model is developed with the express aim of testing the relative impact of reaction rates and branching ratios that have been the subject of much debate over recent years. The importance of an accurate high-resolution solar spectrum is demonstrated. Once good agreement is reached between satellite data and the 1D chemistry model, the model is extended to 3D and implemented in the CMAT GCM. Validation of zonal mean NO profiles calculated by CMAT under different solar and auroral conditions is performed by comparison to recent satellite data.

The following sections will describe satellite observations of NO, an empirical model of NO in the lower thermosphere based on recent observations, and the main production and loss mechanisms for NO. A description of the 1D photochemical model follows, along with results derived using high and low solar and auroral activity conditions. The relative impact of solar energy deposition, auroral energy inputs, key reaction rates and branching ratios are discussed. 3D zonal mean NO profiles calculated by the CMAT model before and after implementation of the 1D photochemical model are presented and compared to recent satellite data. Finally, CMAT calculated NO density profiles produced under differing solar and geomagnetic activity conditions are validated by comparison with an empirical model.

4.2 SATELLITE OBSERVATIONS OF NITRIC OXIDE

Satellite observations of thermospheric NO started in 1968 when the Orbiting Geophysical Observatory (OGO-4) revealed that high latitude NO densities were greater than those observed at low latitudes (*Rusch and Barth* [1975]). Observations made a short time after by the Atmospheric Explorer satellites (*Barth* [1973]) were limited in latitudinal coverage and time. The Solar Mesosphere Explorer (SME) (*Barth* [1992]) measured NO at both high and low latitudes for over 4 years. This data showed that NO concentrations at low latitudes vary with both the 27-day solar rotation and 11-year solar cycle (*Barth et al.* [1988]). A relationship between auroral activity and high latitude NO concentrations was also established (*Barth* [1990], *Barth* [1992]).

The Halogen Occultation Experiment (HALOE) (*Russell III et al.* [1993]), onboard the Upper Atmosphere Research Satellite (UARS), uses solar occultation to measure vertical profiles of various middle atmosphere chemical constituents including NO. The abundance of species is inferred by measuring the attenuation of sunlight by atmospheric gases during sunrise and sunset. This long running data set, started in 1991, covers altitudes from about 10km to 130km over a seasonally varying latitude range. Data from HALOE has been used to gain insight into the impact of large scale tidally driven winds on the spatial and temporal structure of NO (*Marsh and Russell III* [2000], *Marsh and Roble* [2002]). A major drawback arising from the measurement technique is that HALOE only measures a small latitude band each day and takes about 40 days to cover the full range of latitudes (about 120°). Being highly sensitive to changes in auroral and solar energy deposition, NO abundance can change dramatically over this period. HALOE also never samples latitudes above 55° in winter where the largest concentrations of NO are expected to occur.

During the northern spring equinox of 1992, the Atmospheric Trace Molecule Spectroscopy / Atmospheric Laboratory for Applications and Science (ATMOS/ATLAS1) obtained 5.3μm solar occultation spectra data, which yielded peak NO densities of around 10^{14} molecules m^{-3} at mid latitudes (*Krishna Kumar et al.* [1995]). *Swaminathan* [1998] noted that these were in good agreement with NO measurements from HALOE.

Between March 1998 and September 2000, global measurements of NO in the thermosphere were obtained by the Student Nitric Oxide Explorer satellite (SNOE). Simultaneous measurements of solar flux in the soft X-ray (2-10nm) and EUV (10-31nm) wavelength ranges were also taken. SNOE travelled in a near-polar, circular, sun-synchronous orbit and derived NO densities from measurements of fluorescent scattering of sunlight by NO. No observations are made in polar night. Descriptions of the satellite, instrumentation and observation technique can be found in *Solomon et al.* [1996] and *Merkel et al.* [2001]. Details of instrument calibration uncertainties can be found in *Barth et al.* [2003]. Zonally averaged data is available on a 5° latitude grid between ± 80°, at 3.3km height intervals from 97 to 150km altitude. All observations

are within 30 minutes of 10:45 hours LT. This is currently the most spatially comprehensive dataset of NO abundance in the MLT region. While this data is useful for numerical model comparisons, the limited local time coverage means diurnal variations cannot be studied. The high degree of variability present in data sets also makes it hard to identify general trends. It is therefore more productive to use the data set as a whole when comparing to numerical simulations and not rely on data from specific dates. This was made easier by the development of the Nitric Oxide Empirical Model (NOEM) (*Marsh et al.* [2004]) which will be described in the following section.

Observations have revealed that the peak in thermospheric NO concentration typically occurs between 100 and 115km altitude. At low latitudes, concentrations of around 1x 10^{14} molecules m^{-3} are typical, but values can range from half to double that value. NO abundance usually peaks between ±60° and ±80° where values of 2 x 10^{14} molecules m^{-3} are typical. Concentrations of two or three times this value are however not unusual after periods of high auroral activity. The altitude of the peak may also change as high-energy auroral particles penetrate deep into the atmosphere.

4.3 THE NITRIC OXIDE EMPIRICAL MODEL

Marsh et al. [2004] presented a three-dimensional Nitric Oxide Empirical Model (NOEM) that is based on eigenanalysis of SNOE satellite observations. The daily NO data set is represented as a time mean plus the sum of orthogonal functions of space multiplied by time-varying coefficients. The functions, described as empirical orthogonal functions (EOFs) represent the spatial variability in the data set and are functions of geomagnetic latitude and height (y and z). EOFs are ordered by the amount of variance they capture from the original data set, the first capturing the most. The time-varying coefficients, referred to as principal components, describe how this spatial variability varies in time.

Marsh et al. showed that the first three modes of variability identified by the EOFs are associated with variations in three geophysical parameters, namely geomagnetic activity, solar declination angle and solar activity. Using just these 3 EOFs, nearly 80% of the variance in NO density could be explained. Strong correlations between these

geophysical parameters and principal components mean that derived polynomial fits (f_i) of K_p geomagnetic index, solar declination angle, δ, and F10.7 cm flux can be used in place of the principal component time series. The SNOE data set is thus represented as

$$NO(y,z,t) = \overline{NO}(y,z) + f_1(K_p) \cdot E_1 + f_2(\delta) \cdot E_2 + f_3(F10.7) \cdot E_3 \qquad 4.1$$

where $\overline{NO}$ is the time mean NO distribution and E_i are empirical orthogonal functions. The model therefore parameterises the spatial distribution of NO in terms of the planetary magnetic index K_p, day of year, and the F10.7 cm radio flux.

Maximum correlations between the K_p and F10.7 cm flux proxies, and principal component time series 1 and 3, occur with a lag of 1 day. If a comparison between the empirical model and observational data from a specific date is required, geophysical parameters from the previous day must therefore be used as input to the model. This delay in response to changes in solar forcing was also highlighted by *Solomon et al.* [1999] and *Bailey et al.* [2002] and is related to the chemical lifetime of NO in the lower thermosphere.

This empirical model is useful in providing validation of thermospheric NO densities in 3D GCMs such as CMAT as it is able to recreate general trends in the data. Comparisons to this model are preferable to direct comparisons with individual SNOE datasets which can show a large degree of variability. An example of this is seen in Figure 4.11 and Figure 4.12 that show zonal mean NO profiles from days 73 of 1998 and 83 of 2000, taken from the SNOE database. The geomagnetic activity level, as described by K_p index, was the same on both days preceding these dates. As such one might expect the high latitude NO density profiles to be similar. There is however, a factor of 3.1 difference in the magnitude of the high altitude peak in the northern hemisphere, and a factor of 2.3 difference in the south.

In the case of 3D GCMs, we are more interested in reproducing trends and repeatable patterns in atmospheric properties, than individual sets of conditions specific to

individual dates. The aim here is to develop a GCM that can recreate general trends in NO concentration to within a reasonable degree of accuracy, rather than trying to reproduce complex sets of conditions that exist at any one point in time. The NOEM model will therefore be used to provide constraints on NO abundances calculated by both the 1-dimensional photochemical model and CMAT GCM.

4.4 MODELLING STUDIES OF NITRIC OXIDE

Until very recently, photochemical models of the thermosphere have underestimated the magnitude of Nitric Oxide in the MLT region in comparison to satellite and rocket observations. Numerous sensitivity studies have been carried out in order to identify possible sources of the deficit. The importance of solar soft x-rays was first examined by *Siskind et al.* [1990] who added solar fluxes from the 1.8 to 5nm wavelength range to the photochemical model of *Cleary* [1986]. Fluxes 50 times that of the SC21REFW solar reference spectrum (*Hinteregger et al.* [1981]) were needed in order to reproduce the NO density levels observed by rocket experiments. The net effect of increases in soft x-ray flux is to increase the rate of ionisation of N_2. This leads to an increase in production of excited nitrogen atoms ($N(^2D)$) that can react with oxygen to form NO. Soft x-rays contribute relatively little to the direct photoionisation rate but dominate the production of photoelectrons. *Siskind et al.* [1990] used the photoelectron model of *Strickland and Meier* [1982] to demonstrate that, at an altitude of 105km, these photoelectrons produce ten times more ionisation than direct photoionisation. Thus the ratio of the photoelectron impact rate to photoionisation rate was identified as an important component in the modelling of NO.

Barth [1992] used an updated version of the *Cleary* [1986] model, including an enhanced soft x-ray source, to determine the sensitivity of peak NO density to changes in key reaction rates and branching ratios. The most substantial change in the calculated peak NO density occurred when a change was made to the branching ratio for the dissociative recombination of NO^+. The branching ratio for electron impact dissociation of molecular nitrogen was also seen to have a large impact on NO abundance. Uncertainties in the rates for the reaction between excited nitrogen atoms and molecular oxygen, deactivation of excited atomic nitrogen by atomic oxygen, the

photodissociation of NO and the reaction between ionised molecular nitrogen and atomic oxygen all impacted the calculated NO densities. *Barth* [1992] also reported that increases in the flux of auroral electrons lead to a non-linear increase in calculated NO density. While electron precipitation is the dominant process in creation of high latitude NO around 110km, the relationship is complicated by changes in Joule heating associated with geomagnetic activity. Increased heating can result in an increase in NO production above 140km via the temperature sensitive reaction between ground state atomic nitrogen and molecular oxygen. This excess NO may diffuse downward to increase NO abundance at lower altitudes (*Siskind et al.* [1989a, b], *Barth* [1992]).

The response of NO to changes in solar radiation were also investigated by *Fuller-Rowell* [1993]. The globally averaged 1-dimensional model of the thermosphere and upper mesosphere used, was able to reproduce a factor of 7 change in low latitude peak NO density over a solar cycle, similar to changes seen in SME observations (*Barth* [1988]). The study reinforced the importance of soft x-ray flux and the dependence of NO densities on the temperature of the background atmosphere. The inclusion of radiation in the 0.1 to 1.8nm wavelength range was shown to have little impact on the production of NO. The calculated absolute magnitudes of NO were somewhat lower than the SME observations reported by *Barth* [1988] and considerably smaller than the rocket measurements discussed by *Siskind et al.* [1990] and *Siskind and Rusch* [1992]. Although the flux from 1.8 to 5nm was not scaled, as was done by *Siskind et al.* [1990], the need for increased flux in this range was acknowledged if calculated NO densities were to match the high values observed by rockets. It is interesting to note that reprocessed SME data, presented by *Barth* [1996], showed a factor of 2-3 increase in the NO densities reported by *Barth* [1988], bringing them in line with measurements from rockets, HALOE and ATLAS1 (*Swaminathan et al.* [1998]). The discrepancy between model and data thus appears to be greater than initially thought.

Siskind et al. [1995] updated and enhanced the model used in the *Siskind et al.* [1990] study by including solar radiation between 1.1 and 1.9nm, ground state atomic nitrogen production from electron impact dissociative ionisation of N_2, and Auger ionisation. Auger ionisation occurs when a photon has sufficient energy to remove a

K-shell electron from the atom or molecule with which it collides. Once the inner shell electron is released, the vacancy in the K-shell is immediately filled by an outer shell electron. The energy released in the latter transition goes not into radiation, but is released with another outer shell electron. This process results in dissociation of the N_2, molecule and contributes significantly to the total amount of photoionisation. Although individually none of the mechanisms employed in the updated model had a dramatic impact, the cumulative effect was to reduce the size of the soft x-ray scaling factor used in the previous study, from 50 to 25. Note however, that in order to gain good agreement between the model and solar maximum rocket measurements, the flux of radiation in the 1.1 to 1.9nm wavelength band was increased by a factor of 150 over solar minimum values based on the report of *Donnelly and Pope* [1973].

The need for accurate solar soft x-ray fluxes in NO modelling was met in the study of *Swaminathan et al.* [1998], who developed a detailed 1-dimensional photochemical model that incorporated soft x-ray solar fluxes from 1.1 to 5nm, measured by the soft X ray telescope (SXT) on the YOHKOH satellite. The model failed to generate enough NO in comparison to ATLAS1 measurements, suggesting an alternative source for the discrepancy was still outstanding. By updating the NO^+ dissociative recombination branching ratio for $N(^2D)$ from 0.76 to 0.85 (*Vejby-Christensen et al.* [1998]) and replacing the 2-10nm fluxes with empirical data based on SNOE measurements, *Swaminathan et al.* [2001] were able to largely remove the discrepancy between low latitude modelled NO densities and those observed by ATLAS1. YOHKOH data was used for flux at wavelengths below 2nm. A high latitude source of NO was also added, based on the estimate for the year 1998 derived from SNOE data by *Barth et al.* [1999b].

The most recent and seemingly most accurate 1-dimensional photochemical model of NO to date was first presented by *Barth et al.* [1999b], then again in more detail by *Bailey et al.* [2002]. This model is an extension of that used by *Barth* [1992] that includes updated reaction rates and branching ratios. Energetic electron fluxes are calculated numerically using the glow model (*Solomon et al.* [1988], *Solomon and Abreu* [1989]), which includes energetic electron transport, auger ionisation and the

full solar spectrum down to 1.8nm. This removes the need for approximations in photoelectron calculations and for parameterised auroral ionisation rates. By varying the flux of solar soft X-ray flux in three wavelength bands, the importance of flux in the 2-7nm range was once again established. Primary spectra of energetic electrons with given characteristic energies were assumed in order to simulate the relationship between auroral energy input and the altitude of the NO peak. Spectra of higher characteristic energies were shown to produce NO at increasingly lower altitudes. They also showed that the diurnal variability of solar flux and the background atmosphere must be considered to realistically calculate NO abundances, the number of day light hours being essential in determining loss of NO through photodissociation. As such, steady state calculations are not appropriate.

Direct comparisons between the *Bailey et al.* [2002] model and SNOE data were presented by *Barth and Bailey* [2004], who used daily values of 2-7nm flux measured by SNOE as model input. No auroral energy inputs were considered. At latitudes between ±30°, a good match between measured and modelled NO densities was achieved. At higher latitudes, the correlation between data sets decreased. *Barth and Bailey* suggested that the mismatch at latitudes poleward of ±30° was due to aurorally produced NO being transported towards the equator. This study highlights the need for 3D modelling of NO in order that the temporal and spatial scales of transport can be fully explored.

4.5 PRODUCTION AND LOSS OF NITRIC OXIDE

Nitric Oxide in the lower thermosphere is predominantly produced via the chemical reaction between excited atomic nitrogen and molecular oxygen

$$N(^2D) + O_2 \rightarrow NO + O \qquad (1)$$

This is the primary production mechanism at peak altitudes (around 110km). In order for this reaction to proceed rapidly, the nitrogen atoms must have excess energy. Several paths exist for the production of excited nitrogen atoms, the most significant being dissociative recombination reactions that occur with ambient electrons

$$NO^+ + e \rightarrow N(^2D,^4S) + O \quad (2)$$

$$N_2^+ + e \rightarrow 2N(^2D,^4S) \quad (3)$$

Both of these reactions are temperature dependent. Reactions with energetic electrons (e^*) (>10eV) are also important for creation of N(^{2}D) either directly via the reaction

$$N_2 + e^*_{a,p} \rightarrow 2N(^2D,^4S) + e^* \quad (4)$$

where the strong N_2 bond is broken, or through a two step process

$$N_2 + e^*_{a,p} \rightarrow N_2^+ + e + e^*, \text{ followed by} \quad (5)$$

$$N_2^+ + O \rightarrow NO^+ + N(^2D) \quad (6)$$

At low latitudes, high-energy electrons are termed '*photoelectrons*' (e^*_p) and are the product of photoionisation reactions between high-energy photons and neutrals. At high latitudes, they are precipitating auroral electrons (e^*_a). Secondary electrons from processes such as auroral particle bombardment also make up a substantial proportion of the energetic electron population. It is interesting to note that the NO^+ produced in reaction (6) also produces N(^{2}D) through reaction (2).

At altitudes above 120km where the ambient temperature generally exceeds 400K, the reaction between ground state atomic nitrogen, N(^{4}S), and molecular oxygen becomes the dominant source of NO

$$N(^4S) + O_2 \rightarrow NO + O \quad (7)$$

This reaction is slow and very sensitive to temperature.

Two major loss processes for NO occur through reactions with ground state nitrogen, and O_2^+

$$NO + N(^4S) \rightarrow N_2 + O \quad (8)$$

$$NO + O_2^+ \rightarrow NO^+ + O_2 \quad (9)$$

During day light hours, photodissociation of NO by solar UV irradiance is also a significant loss mechanism

$$NO + h\upsilon \rightarrow N(^4S) + O \quad (10)$$

Note that while reactions 9 and 10 destroy NO, they recycle odd nitrogen. If the ground state nitrogen produced in reaction 10 goes on to destroy more NO through reaction 8, the effectiveness of photodissociation as a loss mechanism is increased. Only reaction 8 truly destroys Odd-N. After sunset, although NO is no longer destroyed by solar UV, its abundance is reduced. This is because, with the exception of auroral production, no excited state nitrogen is being created and thus NO production is greatly reduced. Any remaining N(^{4}S) and O_2^+ will react with NO to destroy it through reactions 8 and 9, but in the absence of sunlight, the densities of N(^{4}S) and O_2^+ quickly diminish. Once the loss processes are removed, NO may be transported to lower altitudes and latitudes, especially from auroral regions where NO production continues.

It is clear that the relative abundance of excited and ground state nitrogen is a key factor in both the production and loss of NO. The flux and energy of solar irradiance are fundamental in determining the abundance of these species and of O_2^+. Solar UV can also directly destroy NO through photodissociation. Changes in thermospheric heating can impact key reaction rates, again resulting in changes in NO density. Note that, through consideration of the chemical and diffusive lifetimes of NO *Bailey et al.* [2002] concluded that the abundance of NO at any one time is determined by the level of solar and auroral energy deposition over the previous day.

Although the key processes for the creation and loss of NO have been outlined above, many other reactions must be considered in order to represent the full picture. A table of the reactions calculated in CMAT can be found in Table 3.2.

4.6 A ONE DIMENSIONAL MODEL OF NITRIC OXIDE

The 1D model calculates constituent profiles for a given set of geophysical conditions, at a set location on the globe. The vertical range covers 63 pressure levels at 1/3 scale height from a lower boundary pressure of 10mb at approximately 30km altitude. The upper boundary of the model ranges from 250 to 600km depending on solar activity. The density of NO, NO_2, $N(^4S)$, $N(^2D)$ and NO_x are calculated by solving the one-dimensional mass continuity equation, including vertical transport

$$\frac{\partial \chi_i}{\partial t} = \frac{1}{\rho} m_i S_i + \frac{1}{\rho}\frac{\partial}{\partial P}((D_{EDDY} + D_{MOLEC})n\frac{\partial}{\partial P}(m\chi_i)) \qquad 4.2$$

where $\chi_i = n_i m_i/\rho$ is the mass mixing ratio of species *i*, S_i denotes chemical sources and sinks, m_i molecular mass, n_i number density, ρ mass density, D_{MOLEC} the molecular diffusion coefficient, and D_{EDDY} the eddy diffusion coefficient. The first right hand term corresponds to chemical production and loss, the second and third to turbulent and molecular diffusion respectively. The production and loss of NO are principally due to the processes described in the previous section. Molecular and eddy diffusion coefficients are as described in section 2.8.1 and 2.7.2 respectively.

A number of other species in the model undergo chemical processes much more quickly than diffusive processes meaning photochemical equilibrium can be assumed. Densities of these species, namely N_2^+, N^+, O_2^+, O^+ and NO^+ are calculated by considering the chemical production and loss terms alone.

The start up profiles of all species are taken from CMAT, along with profiles of O_3, N_2O, OH, $O(^1D)$ and HO_2. Minor species are calculated with an integration period of 5 minutes. Major constituent profiles, profiles of OH, $O(^1D)$ and HO_2, neutral, ion and electron temperatures are all taken from CMAT and updated every time step such that the daily variation in solar zenith angle is represented. This process is repeated until the model reaches convergence, taken as the point when diurnal variability in NO concentration is repeated from day to day.

The solar energy input to the model ranges from 0.2 to 180nm, covering the principal ionising region of the solar spectrum. For wavelengths less that 1.8nm, flux of X-rays are taken from the GLOW model (*Solomon et al.* [1989], *Solomon* [2001]). Fluxes longward of 1.8nm are taken from the SOLAR2000 empirical model (*Tobiska et al.* [2000]). Photoionisation and absorption cross sections are based on *Henke* [1993] for wavelengths between 0.2 and 6nm, *Fennelly and Torr* [1992] (*Stan Solomon and Liying Qian*: private communication) between 6nm and 105nm, and multiple sources for higher wavelengths (*Rodney Viereck*: private communication). Ionisation and dissociation rates are calculated as described in section 1.3.8. Chapman grazing incidence functions are calculated for each absorbing species to account for curvature of the Earth at large solar zenith angles (*Smith and Smith* [1972]).

As previously mentioned, one of the products of primary photon ionisation is a photoelectron. If sufficiently energetic, these photoelectrons can go on to cause multiple secondary ionisations, greatly enhancing the ionisation rate of N_2 and thus the production of NO. Photoelectron ionisation as calculated by the code of *Strickland and Meier* [1982], described by *Fuller-Rowell* [1993] is included in CMAT and the 1D model. Comparison of the low latitude ionisation rates produced by the 1D model with those calculated by *Bailey et al.* [2002] using the glow model, show a shortfall in the 1D model ionisation rates for all major species. This may be due in part to the photon energy spectra used in the *Strickland and Meier* model which calculated the photoelectron distribution for energies up to 450eV. Higher energy photons and primary photoelectrons were not accounted for. Another possible cause for the discrepancy is that vertical profiles of the ratio of photoelectron ionisation to primary ionisation of the major species were only determined down to 100km. Below this altitude, the 1D model uses extrapolated values of O and O_2 ratios, and an estimate of the N_2 ratio based on the available energies of photons that would penetrate to a given altitude. Finally, *Siskind et al.* [1995] showed that the inclusion of Auger ionisation led to a significant increase in the N_2 ionisation rate. This effect is not accounted for in CMAT or the 1D model. In order to account for the effects of Auger ionisation and high-energy photons/photoelectrons, the ratios of ionisation due to photoelectrons

compared to primary ionisation have been enhanced in line with those given by *Bailey et al.* [2002].

Figure 4.1 (*top*) shows the rates of O, O_2 and N_2 ionisation and N_2 dissociation at the equator due to solar irradiance produced photoelectrons as calculated by the 1D model. Background conditions are as for day 266 in 1999. The local time is 11:00AM and the solar soft flux (2-7nm) is set at $1mWm^{-2}$.

Figure 4.1 (*bottom*) shows similar plots as calculated by the glow model under the same set of conditions (*Bailey et al.* [2002]). The morphology and magnitude of both sets of rates is in reasonable agreement, with the maxima occurring at around 110km altitude.

Auroral energy deposition is derived from measurements of precipitating auroral particles from the TIROS/National Oceanic and Atmospheric Administration (NOAA) Space Environment Monitor (SEM) as described by *Fuller-Rowell and Evans* [1987] and *Codrescu et al* [1997]. For particle energies between 300eV and 30keV, the shape of the energy spectrum is derived from the 'characteristic energy', determined from measurements of the maximum sensor response occurring during a sweep from 300eV to 30keV. For a Maxwellian particle spectrum this characteristic energy would be the average energy of the distribution. These energies have been weighted with the corresponding energy flux measurements to obtain statistical maps of energy influx. All particles in the 30eV to 20keV energy range are assumed to be electrons. The effect of particles with energies ranging from 30keV to 2.5MeV is included through use of precipitation patterns derived by *Codrescu et al.* [1997].

Fuller-Rowell and Evans [1987] used a simple expression to calculate the total ionisation rate resulting from the precipitation patterns, considering the characteristic energy and incident flux of particles and assuming the model atmosphere temperature and composition profile of *Jones and Rees* [1973]. Ionisation rates of N_2, O_2 and O are then determined from relationships given by *Jones and Rees* that take into account differences in ionisation cross-sections for each species. These rates are applied in both CMAT and the 1D model. For a given set of characteristic energies and an incident

energy flux of $1mWm^{-2}$, the altitude profiles of N_2 ionisation rate produced match reasonably well with those calculated using the glow model (*Bailey et al.* [2002]) which performs a full calculation of energetic electron flux including electron transport. This suggests that the simple calculation performed by *Fuller-Rowell and Evans* is sufficient to calculate a reasonable N_2 ionisation rate profile.

Figure 4.2 shows the total ionisation rate at 120km during conditions of low (K_p 0 to 1), medium (K_p 2 to 3) and high (K_p 4 to 5) geomagnetic activity, calculated using precipitating electron data from the Particle Environment Monitor (PEM) onboard UARS (*Ridley* [1999]). Each plot has an outer edge at 50° latitude and is centred on the magnetic pole. The right and left hand panels show the southern and northern hemispheres respectively. As activity increases, the equatorward edge of the auroral oval moves to lower latitudes and the ionisation rate increases. For medium to high geomagnetic activity, the total ionisation rate at 120km calculated using characteristic electron energies and fluxes from the TIROS precipitation data is in reasonable agreement with those presented by *Ridley*. At low geomagnetic activity however, the ionisation rate at 120km is lower than that derived from PEM data. The total ionisation rate due to particle precipitation has been increased by a factor of 4 in the 1D model for K_p index values below 2. This factor was chosen, as the corresponding profiles of high latitude NO are in much better agreement with NO profiles from the NOEM model than those created with the low K_p TIROS data. This will be discussed further in the next section. This increase leads to an ionisation rate that is somewhat higher than the rate derived from PEM data at 120km. The maximum in the enhanced TIROS derived rate is however largely within 20° of the geomagnetic pole where PEM derived rates are not available. Figure 4.3 shows ionisation rate maps for the northern and southern hemispheres for conditions of low, medium and high geomagnetic activity as derived from the TIROS data and used in CMAT and the 1D model. Each plot has an outer edge at 50° latitude and is centred on the geographic pole. The position of the magnetic pole being indicated by a red diamond symbol on each plot. The right and left hand panels show the southern and northern hemispheres respectively. The low activity rates shown have been enhanced by a factor of 4 over those calculated by *Fuller-Rowell and Evans*.

Possible reasons for the apparent underestimation of ionisation rates in the calculations of *Fuller-Rowell and Evans* include the limited coverage of the TIROS measurements, which only sampled a small portion of the auroral oval at any one time. The data returned from the SEM instrument was reasonably coarse energy spectra data from limited local solar times. At low activity, the number of data points in each bin is low (*Codrescu et al.* [1997]), thus accuracy is reduced. Another possible reason for error is that all keV particles are assumed to be electrons. *Galand et al.* [1999] showed that this is not a good assumption as protons can represent an important source of ionisation.

4.7 1D MODEL LOW LATITUDE NITRIC OXIDE DENSITIES

Figure 4.4 shows Nitric Oxide densities at the equator as calculated by the 1D model, as given by the Nitric Oxide Empirical Model (NOEM) and as given by the Student Nitric Oxide Explorer (SNOE) satellite. Geophysical conditions are appropriate to day 72 of 1998 on which the F10.7 was 104. Using F10.7 as a proxy for solar activity, this day is one of the lowest activity days within the SNOE dataset and as such has been chosen to represent low solar activity conditions. By considering the lifetimes of both chemical destruction and diffusive transport of an NO molecule, *Bailey et al.* [2002] deduced that the NO abundance at any one time is indicative of the solar energy input over the past day. The SNOE data shown is therefore from day 73 of 1998 and is representative of the solar and auroral conditions from the previous day. The morphology of all 3 profiles are in good agreement with the peak in NO density occurring close to 110km where solar soft X-rays are absorbed. NO densities drop off sharply above and below this altitude. Maxima of 8.3, 7.1 and 6.5 $x10^{13}$ molecules m^{-3} are produced by the 1D model, NOEM and SNOE respectively.

As previously mentioned, NO densities measured by SNOE on individual dates show a high degree of variability meaning comparisons of specific days can be misleading. An appreciation of the upper and lower limits of the peak NO density for the conditions simulated here can be gained by viewing NO densities recorded by SNOE on dates with equivalent solar conditions. Such an analysis reveals minimum and maximum NO peak densities of 5.8 and 7.8 $x10^{13}$ molecules m^{-3} present in data sets recorded one day

after an F10.7 of 103 has been measured. The NO maximum produced by NOEM is therefore within the limits of the maximum measured by SNOE, while the 1D model maximum is slightly larger.

Figure 4.5 shows equatorial NO densities given by the 1D model, NOEM and SNOE for conditions of high solar activity (F10.7 = 232). Geophysical conditions are appropriate to day 82 of 2000. Again the morphologies of the profiles are in good agreement, with the peak NO density occurring between 105 and 110km altitude. The NO density given by the 1D model and the SNOE data is slightly higher that that given by the NOEM at all altitudes. The peak values of 1.8 (1D model), 1.2 (NOEM) and 1.3 (SNOE) $x10^{14}$ molecules m^{-3} are in reasonable agreement, the upper and lower limits of SNOE measurements being 2.5 and 1.2 $x10^{14}$ molecules m^{-3} for this level of solar activity. It is clear from comparison of Figure 4.4 and Figure 4.5 that NO abundance in the middle atmosphere is strongly influenced by solar activity, increasing at all altitudes with increasing F10.7. The largest changes in NO density are seen at and below the NO peak where progressively higher energy solar photons are absorbed.

4.8 1D MODEL HIGH LATITUDE NITRIC OXIDE DENSITIES

NO density profiles at 65°N are shown in Figure 4.6 as given by the 1D model, NOEM and SNOE. Conditions are appropriate to day 79 of 1998 when geomagnetic activity was low (K_p = 1°). Even at low activity, the NO abundance at high latitude is clearly greater than at low latitude. This is due to the increase in N_2 ionisation arising from auroral energy input. In each plot, the NO peak occurs close to 110km altitude, with values of 1.6, 2.3 and 2.4 $x10^{14}$ molecules m^{-3} given by the 1D model, NOEM and SNOE data respectively. By considering dates throughout the SNOE data set where equivalent levels of auroral activity occurred, lower and upper limits for the peak NO density can be set at 1.1 and 2.5 $x10^{14}$ molecules m^{-3} respectively for a K_p of 1°. NO densities produced by the 1D model are slightly lower than those given by NOEM and SNOE at all altitudes in the auroral zone. This is most likely due to inaccuracies in the auroral energy input for low activity conditions as discussed above.

Results for high geomagnetic activity are shown in Figure 4.7. Conditions are appropriate to day 240 for 1998 when the average K_p was 4^+. All the profiles show a considerable increase in NO abundance over those produced in conditions of low geomagnetic activity. Maxima of 4.9, 2.9 and 5.5 $x10^{14}$ molecules m^{-3} occur at 113, 105 and 104 km altitude for the 1D model, NOEM and SNOE respectively. Maximum and minimum peak NO densities measured by SNOE on dates with equivalent levels of geomagnetic activity are 5.6 and 2.0 $x10^{14}$ molecules m^{-3} respectively. The peak NO density produced by the 1D model occurs approximately 9km higher than expected and high altitude densities are somewhat greater than those produced by NOEM and SNOE. This is again attributed to inaccuracies in the representation of auroral energy input at high latitudes. The discrepancy between the height of the modelled and data NO peak suggests that the flux of deeply penetrating high-energy particles is underestimated in the model. The flux of lower energy particles may be overestimated, resulting in the enhanced NO density above the peak.

4.9 MODEL SENSITIVITY TO SOLAR SPECTRUM

It is now well accepted that the flux of solar energy in the 2 to 7nm wavelength range plays a key role in the production of NO in the middle atmosphere. *Bailey et al.* [2002] showed that inclusion of $1mWm^{-2}$ integrated energy flux in the 2-7nm band had a dramatic effect on NO densities calculated by their 1D photochemical model. Varying the flux of energy in this band led to changes in the magnitude of the NO density peak. Empirical solar flux data were however not used in the *Bailey et al.* study and the question of how simulated NO densities vary with realistic changes in solar flux input was not addressed. Given the variety of solar reference spectra used in modelling studies, it is useful to gain some appreciation of how NO densities vary with differing reference spectra. Figure 4.8 shows the equatorial NO density calculated by the 1D model under conditions of low (F10.7 = 104) and high (F10.7 = 232) solar activity using solar flux data based on the commonly used *Hinteregger et al.* [1981] spectrum and using fluxes from the SOLAR2000 empirical model (*Tobiska et al.* [2000]). Comparison of the NO density profile produced using solar flux data from each spectrum reveals a large increase in NO density when using the SOLAR2000 data. The peak density increases from 4.6 to 8.3 $x10^{13}$ molecules m^{-3} and moves down in altitude

from 114 to 110km. While the largest difference in NO density is seen at the peak, significantly more NO is present at all altitudes below about 130km when using flux data from the SOLAR2000 model.

Under conditions of high solar activity the difference in NO abundance is more dramatic with the peak density changing nearly a factor of 4 from 0.53 to 1.8 $x10^{14}$ molecules m^{-3}. Again there is a slight downward shift in the altitude of the peak when using the SOLAR2000 flux data. A large difference in NO concentration is evident at all altitudes below about 150km. Comparison with the NO profiles produced by NOEM and SNOE data in Figure 4.4 and Figure 4.5 shows that the densities calculated when using the SOLAR2000 fluxes are in much better agreement with satellite data than when the *Hinteregger et al.* fluxes are used.

In the model of *Bailey et al.* [2002], a factor of 2 increase in the 2-7nm solar soft X-ray flux led to approximately the same factor increase in peak NO density at the equator. This would suggest a roughly linear relationship between the two. With this in mind, the results here imply that the soft X-ray fluxes given by the *Hinteregger et al.* solar reference spectrum are approximately a factor of 2 to 4 too low. This is in line with the findings of *Solomon et al.* [2001] who proposed an increase in the *Hinteregger et al.* [1981] fluxes of around a factor of 4 for wavelengths below approximately 25nm. This was based on improved photoelectron spectra and electron density profiles calculated using newly measured solar soft X-ray fluxes from the solar X-ray photometer (SXP) onboard SNOE. The case for an increase was also stated by *Richards et al.* [1994] and by *Richards and Torr* [1984] who suggested a factor of approximately 3 increase in the *Hinteregger et al.* solar fluxes below 25nm.

4.10 MODEL SENSITIVITY TO KEY BRANCHING RATIOS

There is still some uncertainty over several of the key reaction rates and branching ratios used in 1D photochemical models of NO. The photochemical scheme used in the 1D model developed here was based on the original CMAT scheme. Many of the reaction rates and branching ratios required updating to bring them in line with current literature. A tuning process was then performed to find the set of rates that led to the

best fit between model and data under conditions of low and high solar and auroral activity.

Of particular importance is the branching ratio of electron impact NO^+ dissociative recombination (Reaction 2 above) for $N(^2D)$ which has been updated from 0.76 (*Kley et al.* [1977]) to 0.85 following *Swaminathan et al.* [2001].This resulted in a 77% increase in the equatorial peak NO density from 4.7 to 8.3 $x10^{13}$ molecules m^{-3} at 110km, during low solar activity. At 65° latitude the change in branching ratio resulted in a peak NO increase of 150%. NO densities at all altitudes below 140km increased by varying degrees following the rate change.

Another important branching ratio is that of electron impact dissociation of molecular nitrogen (reaction 4). The sensitivity of the model to this value was tested in the range 0.54 to 0.62 (the range of estimated uncertainty according to *Swaminathan et al.* [1998]). Considering all auroral and solar conditions tested, the best NO profiles resulted from an $N(^2D)$ yield of 0.54. This value is consistent with experiments by *Swaminathan et al.* [2001]. At 65°, this change in ratio results in a decrease of around 67% in peak NO density.

4.11 MODEL SENSITIVITY TO KEY REACTION RATES

The rate coefficient for reaction 8 above has been updated from 3.4 x 10^{-17} m^3s^{-1} to include a temperature dependency in line with *Swaminathan et al.* [1998]. For temperatures greater than 400 K, a rate of 3.25 x 10^{-17} m^3s^{-1} is used. For temperatures below 400 K, a rate of 2.2 x 10^{-17} exp $(160/T_n)$ m^3s^{-1} is used where T_n is the neutral temperature. This resulted in a maximum increase in the NO peak of about 10%. The sensitivity of the model to this rate was tested in line with the limits suggested by *Wennberg et al.* [1994] i.e. (2.2 ± 0.2) x 10^{-17} exp $((160 \pm 50)/T_n)$ but changes to either variable element had little effect on NO densities. Tests were also carried out using a rate of 2.1 x 10^{-17} exp $(100/T_n)$, as used in the 2D model of *Vitt et al.* [2000]. Use of this rate was found to result in an overestimation of NO under all conditions tested.

The rate for the reaction between ground state nitrogen atoms and molecular oxygen (Reaction 7) has been updated from $4.4 \times 10^{-18} \exp(-3220/T_n)$ m^3s^{-1} (*Barth* [1992]) to $1.5 \times 10^{-17} \exp(-3600/T_n)$ m^3s^{-1} as used by *Siskind et al.* [2004]. This leads to an increase in the magnitude of the NO peak of between 22 and 28% at both low and high latitudes. The change in NO at higher altitudes is more dramatic with an increase of over 70% occurring at 150km for high and low latitudes.

One more reaction rate coefficient found to have a significant effect on calculated NO abundance is for the reaction between excited state nitrogen atoms and molecular oxygen (Reaction 1). This rate has been updated from 5.9×10^{-18} m^3s^{-1} to $6.2 \times 10^{-18} \exp(T_n/300)$ m^3s^{-1} after *Duff et al.* [2003]. The result is an increase in peak NO density of around 25%. At altitudes above the peak, the increase in NO is dramatic and reaches over 100% at 150km. A summary of the key reaction rates and branching ratios, and their effect on NO density is given in Table 4.1.

Reaction	**Reaction rate (m^3s^{-1}) or branching ratio (%)**		**Maximum NO increase**
	Previous	**New**	
$NO^+ + e \rightarrow N(^2D,^4S) + O$	0.76 (N^2D)	0.85 (N^2D)	77% at peak (0°) 150% at peak (65° N)
$N_2 + e^* \rightarrow 2N(^2D,^4S) + e^*$	0.75 (N^2D)	0.54 (N^2D)	67% at peak (65° N)
$NO + N(^4S) \rightarrow N_2 + O$	3.4×10^{-17}	$T_n > 400$ K: 3.25×10^{-17} $T_n < 400$ K: $2.2 \times 10^{-17} \exp(160/T_n)$	≈ 10% at peak
$N(^4S) + O_2 \rightarrow NO + O$	$4.4 \times 10^{-18} \exp(-3220/T_n)$	$1.5 \times 10^{-17} \exp(-3600/T_n)$	≈ 22 - 28% at peak > 70% at 150km
$N(^2D) + O_2 \rightarrow NO + O$	5.9×10^{-18}	$6.2 \times 10^{-18} \exp(T_n/300)$	≈ 25% at peak >100% at 150km

Table 4.1 Summary of key reaction rate and branching ratio updates in the 1D photochemical model, and their maximum resultant effect on NO number density.

4.12 APPLICATION IN CMAT

The 1-dimensional photochemical model described above has been extended to calculate chemical production and loss of NO and related species in 3 dimensions. In integrating this model into the CMAT GCM, the effects of horizontal and vertical advection are added to the minor species calculation, resulting in a more comprehensive picture of the global NO distribution.

Figure 4.9 and Figure 4.10 show zonal mean NO densities at 10:30am LT calculated by CMAT before and after the NO photochemical model was added. The model has been run for a period of 40 model days such that it is in a '*steady state*'. Here steady state means the diurnal variability is repeated from day to day. NO densities from the SNOE satellite are also shown for comparison. The densities in Figure 4.9 were calculated using geophysical conditions appropriate to day 266 of 1999 when the F10.7 was 137 and the K_p was 4°. The profiles in Figure 4.10 were calculated using solar and geomagnetic activity conditions appropriate to day 79 of 1998 with an F10.7 of 126 and a K_p of 1°. As discussed earlier, the SNOE data shown is from one day later, as this day is representative of the conditions of the previous day. All plots show the expected latitudinal structure in NO abundance, with enhanced quantities in the high latitude auroral zones and peak densities occurring in the middle atmosphere. The degree of zonal asymmetry, particularly visible in the auroral zones, is related to the offset between geographic and geomagnetic poles. When geographic coordinates are used, as is the case here, the asymmetry is most pronounced in the southern hemisphere where the offset is largest.

The magnitudes of NO abundances calculated by the original CMAT model are clearly deficient at all latitudes. Note the difference in scale between the plots. In Figure 4.9 while the low latitude peak in NO density is positioned around 110km as expected, the magnitude of the peak is a factor of 10 too low. At high latitudes, the location of the NO peak is between 130 and 140km, considerably higher than the peak in the SNOE data which occurs close to 110km. The magnitudes of the high latitude peaks are between 5 and 6 times lower in the old CMAT calculations than those given by SNOE

data. The high latitude peaks are also very localised in the old CMAT calculation, with very little NO equatorward of ±60°. The bottom plot in Figure 4.9 shows that inclusion of an up-to-date photochemical scheme in CMAT has led to the total removal of the large deficiency in calculated NO densities under conditions of moderate solar and high geomagnetic activity. The low latitude peak values are now in line with those given by the SNOE satellite with maxima of around 1 x $10^{14}m^{-3}$ located close to 110km. There is good agreement between both the altitude and magnitude of the high latitude peaks given by SNOE and calculated by the new CMAT. The vertical extent of the high latitude NO is greater in the CMAT calculations than given by SNOE but this is to be expected considering the average nature of the auroral energy inputs used in the model. The latitudinal spread in NO is now in much better agreement with the satellite data with high concentrations present at mid to low latitudes.

In Figure 4.10, the abundance of NO at low latitudes is slightly less than in Figure 4.9 because of the lower level of solar activity (F10.7 = 126 on day 79 1998, vs. F10.7 = 137 on day 266 1999). In all three plots, the equatorial NO number density peak occurs around 110km, and has values of 1.8, 8.0 and 8.0 x 10^{13} m^{-3} as given by the old CMAT code, SNOE and the updated CMAT code respectively. As seen in the previous figure, the large discrepancy between CMAT and SNOE NO densities has been removed by updates to the photochemical scheme. At high latitudes, the NO maxima are considerably lower than those in Figure 4.9 due to the difference in geomagnetic activity levels. The K_p on day 79 of 1998 was 1° indicating very low geomagnetic activity whereas the K_p of 4° on day 266 of 1999 indicates conditions of relatively high geomagnetic activity. The high latitude NO peaks calculated by the old CMAT code are a factor of 11 to 15 less than those given by SNOE over the high latitude peak region. Auroral peak magnitudes calculated by the new CMAT code are in better agreement with the SNOE data with maxima of 1.6 x 10^{14} m^{-3} occurring close to 80° N and 1.3 x 10^{14} m^{-3} near 80° S. The SNOE maxima of 2.4 and 2.0 x 10^{14} m^{-3} occur between 65-70° in the northern hemisphere and around 80° in the south. The location of the northern hemisphere high latitude NO peaks calculated by CMAT are slightly poleward of those given by SNOE and the equatorward spread is less pronounced.

There is good agreement between the vertical extent of the peak NO number density calculated by the new CMAT model and that given by the SNOE satellite.

In order to test the ability of the CMAT model to reasonably recreate NO abundances under different solar and geomagnetic conditions, the model was run under conditions of low and high solar activity and under low and high geomagnetic activity. Conditions were chosen that are appropriate to 4 dates within the SNOE dataset such that comparisons with real data can be made. The NOEM model has been run for the same sets of conditions such that the ability of CMAT to recreate general trends in NO density can be tested. The results from these runs are shown in Figures 4.11 to 4.14.

Zonal mean NO densities at 10:30 LT are shown in Figure 4.11 as calculated by the CMAT model, the NOEM model and as given by the SNOE satellite. Conditions are appropriate to day 72 of 1998 when solar activity was low (F10.7 = 104) and geomagnetic activity was moderate ($K_p = 3^-$). The SNOE data shown is from day 73 of 1998, this being representative of the conditions on the previous day. As expected, there is considerably more NO at high latitudes than at low due to auroral energy sources. The peak NO density is centred close to 110km in all data sets, the NOEM peak being located slightly lower at low and northern latitudes, and the CMAT peak slightly higher at low latitudes. Equatorial maxima of 7.0, 7.2 and 6.9 x 10^{13} m^{-3} are given by the CMAT model, the NOEM model and the SNOE satellite data respectively. There is also good agreement between mid latitude abundances in each dataset. The location and spatial extent of the high latitude peaks are in reasonable agreement, the northern hemisphere peak calculated by CMAT being slightly poleward of those given by NOEM and SNOE, and the vertical extent of the southern peak being slightly higher. The magnitude of the high latitude peaks given by SNOE are considerably higher than those given by CMAT and NOEM, which both reach a maximum of 2.4 x 10^{14} m^{-3} in the northern hemisphere.

The zonal mean NO density profiles shown in Figure 4.12 represent conditions of high solar activity and moderate to high geomagnetic activity. Both NOEM and CMAT were run with an F10.7 of 232 and a K_p of 3^-, in line with the geophysical conditions of

day 82 in the year 2000. The SNOE plot shown is from day 83 of 2000. Low latitude NO densities are considerably higher than those in Figure 4.11 due to the increase in solar activity. The altitude of the NO peak is located close to 110km in all three plots. The low latitude CMAT peak being slightly higher and the high latitude NOEM peak being slightly lower in altitude than those seen in the SNOE data. A comparison of the plots reveals an overestimation of low latitude NO densities calculated by CMAT. Equatorial maxima of approximately 1.9, 1.2 and 1.3 x 10^{14} m^{-3} are present in the CMAT, NOEM and SNOE data respectively. The 2-dimensional models of *Fuller-Rowell and Rees* [1996], *Siskind et al.* [1997] and *Vitt et al.* [2000], also predicted low latitude peak NO concentrations higher than those observed. This may be related to the vertical diffusion coefficient used in CMAT which is at present invariant with latitude. *Siskind et al.* found better agreement between their model and observations when the vertical diffusion coefficient was increased, effectively transporting more NO into the upper mesosphere.

There is good agreement between mid latitude densities in all three plots. The magnitude of the high latitude NO peak densities given by NOEM and CMAT in both the northern and southern hemispheres are in good agreement although the vertical extent of the CMAT peak is greater than that in NOEM. A possible reason for this is the temperature sensitivity of the reaction between ground state atomic nitrogen and molecular oxygen. Small increases in thermospheric temperatures above about 120km can result in excess production of NO. Auroral NO densities in the southern hemisphere section of the SNOE data are lower than those at mid and low latitudes. This highlights the fact that general trends in NO abundance are not always visible in the highly variable SNOE data sets of individual days.

Figure 4.13 shows the zonal mean NO densities calculated by CMAT, NOEM and as give by SNOE under conditions of low geomagnetic activity (K_p = 1°) and moderate solar activity (F107 = 124). These conditions are representative of day 78 in 1998. As in the previous figures, the SNOE data is from the following day, i.e. day 79 of 1998. The altitude of the low latitude NO peak is slightly higher in the CMAT data (around 115km) than in the NOEM and SNOE data where the peak is located around 108km.

The magnitude of the low latitude peak densities are in reasonable agreement in each plot with equatorial maxima of 1.0, 0.8 and 0.8 x 10^{14} m^{-3} given by CMAT, NOEM and SNOE respectively. There is however a visible enhancement in the low latitude NO abundance predicted by CMAT. This local maximum, along with the slight overestimation in peak altitude, may again be related to the vertical diffusion coefficient used.

Mid latitude NO maxima are in reasonable agreement in each plot although the CMAT northern hemisphere densities are slightly lower than those given by NOEM and SNOE. Possible reasons for a deficiency in the calculated mid latitude densities are related to the magnitude and spatial distribution of the auroral energy inputs. The low activity precipitation patterns used in CMAT may be located too far poleward, or may have a limited latitudinal spread resulting in a localised NO peak. The equatorward circulation associated with Joule heating may be underestimated, restricting the horizontal transport of atmospheric constituents. Another related factor is that compressional heating, arising from the interplay between the solar driven poleward circulation and the aurorally driven equatorward circulation, may effect the amount of NO produced through the temperature sensitive reaction between $N(^4S)$ and O_2.

The magnitude and altitude of the high latitude peaks are in good agreement in each data set. CMAT giving peaks of 1.9 and 1.5 x 10^{14} m^{-3} at 110km at latitudes of 82° and -84° in the northern and southern hemispheres respectively. NOEM high latitude peaks of 1.7 and 1.5 x 10^{14} m^{-3} occur just below 110km in the northern and southern hemispheres at latitudes around 70° and –65°. The SNOE high latitude peaks are located close to 70° and -80° at altitudes of 110km in the north and 105km in the south. Both northern and southern hemisphere SNOE NO peaks are just under 2.0 x 10^{14} m^{-3} in magnitude. As in the previous figure, the high latitude NO peaks predicted by CMAT extend slightly further into the thermosphere than those given by NOEM.

Results from CMAT and NOEM model runs under conditions of high auroral activity are shown in Figure 4.14 along with SNOE data from the following day. Conditions for the runs were F107 = 138 and K_p = 4°, representative of day 266 of1999. Overall

agreement between the data sets is good with high latitudes peak densities dominating over low. As expected, the auroral peaks are much greater than in Figure 4.13 where the geomagnetic activity was low. The altitude of the peak is again located close to 110km although the NOEM model predicts the peak to be slightly lower than CMAT or SNOE. The magnitude of the low and mid latitude NO densities are in good agreement in all three cases, as is the latitudinal distribution. The high latitude peaks are all located between 60° and 65° in the northern hemisphere, NOEM giving the largest NO density peak with a magnitude of 2.9 x 10^{14} m^{-3} at 65°, compared to the CMAT and SNOE peaks of 2.5 x 10^{14} m^{-3} and 2.7 x 10^{14} m^{-3} located at 65° and 60° respectively. In the southern hemisphere, the latitudinal spread of the maximum NO density is more dispersed, occurring between -50° and -60° in the CMAT data, between –60° and –65° in the NOEM data and between 75° and 80° in the SNOE data. The magnitude of the southern hemisphere peaks are in good agreement in all 3 plots and range from 2.3 x 10^{14} m^{-3} in both CMAT and NOEM data, to 2.4 x 10^{14} m^{-3} as given by SNOE. Statistical blurring in the precipitation and electric field models used can result in a smoothing out of the local response to auroral activity and a reduction in the peak magnitude compared to instantaneous observations (*Fuller-Rowell et al.* [1994]). The broadened latitudinal NO distribution and reduced peak in the northern hemisphere are thought to be an attribute of this statistical smoothing. As previously discussed, the vertical extent of the high latitude peaks calculated by CMAT is somewhat extended over that given by NOEM and SNOE.

4.13 CONCLUSIONS

A 1-dimensional photochemical model of nitric oxide in the lower thermosphere has been developed that takes into account a total of 37 ion-neutral and neutral-neutral chemical reactions. NO density profiles calculated by the model have been compared to profiles from a recently published empirical NO model (NOEM) under conditions of high and low solar and geomagnetic activity, as defined by the F10.7 cm flux solar activity proxy and the K_p planetary geomagnetic index. While NO profiles calculated by the 1D model under conditions of high solar activity are slightly higher than those

given by the empirical model, there is overall, a reasonable agreement between the two models under all geophysical conditions tested.

Two commonly used solar reference spectra have been implemented in the photochemical model in order to determine the impact of the resultant photolysis rates on low latitude NO densities. The best fit between simulated and observed equatorial NO densities resulted from use of the SOLAR2000 empirical model (*Tobiska et al.* [2000]). Under conditions of high solar activity, use of the SOLAR2000 spectrum led to nearly a factor of 4 increase in the simulated peak NO density over that calculated using fluxes based on the *Hinteregger et al.* [1981] spectrum. This is in line with increases of approximately 3 to 4 that have previously been proposed.

Sensitivity tests to changes in key branching ratios revealed that calculated NO densities are strongly affected by the branching ratios of both electron impact NO^+ dissociative recombination and electron impact dissociation of molecular nitrogen. Significant changes in NO profiles resulted from changes to certain key reaction rate coefficients. Most important were the rates for the reaction between NO and ground state atomic nitrogen, and the reactions between molecular oxygen and both ground and excited state atomic nitrogen.

Inclusion of the 1D photochemical model in the CMAT GCM has led to the total removal of the large deficiency in calculated NO densities that was previously present at all latitudes. The altitude of the calculated NO density peak has also been constrained in line with satellite data. The improved CMAT code is able to reasonably recreate zonal mean NO number density profiles as given by NOEM under conditions of low and high solar and geomagnetic activity. Notable discrepancies between model and observational data arise from the tendency of CMAT to overestimate the equatorial NO peak, to shift the location of the predicted high latitude NO peak poleward by 10 to 15 degrees latitude in both hemispheres during conditions of low auroral activity, and to overestimate the vertical extent of the high latitude peaks. These shortfalls are thought to be related to the vertical diffusion coefficient and statistical auroral energy inputs used in the model.

Despite these small disparities between the empirical and CMAT models, the match between NO abundances is thought to be sufficiently good that CMAT can now be used to study temporal and spatial variation in NO densities under conditions of low and high solar and auroral activity. This will be of particular use in simulating the thermal response of the atmosphere to periods of high geomagnetic activity when excess NO can act to cool the thermosphere. Transport of high latitude thermospheric NO from the upper to lower atmosphere is thought to be one mechanism by which the two areas are coupled. Several studies have suggested that this transport takes place in the cold winter pole where loss of NO through photodissociation is low, during periods of advective descent (e.g. *Callis et al.* [1996], *Callis et al.* [1998]). This additional source of NO_x in the stratosphere can act to catalytically destroy ozone (*Randall et al.* [2001]). CMAT modelling studies may play some part in addressing the nature and extent of this process.

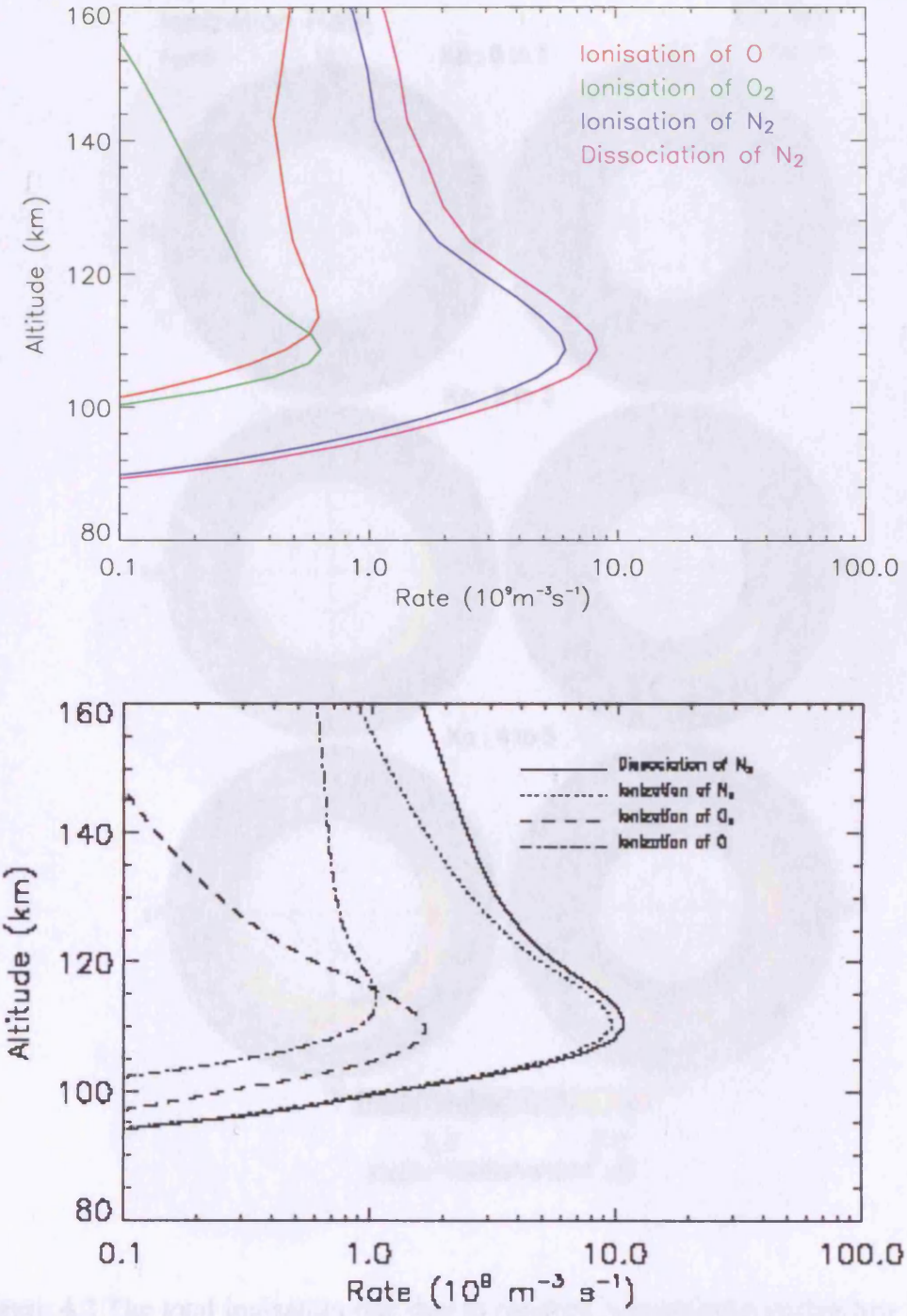


Figure 4.1 Ionisation rates of N_2, O_2, O, and the dissociation rate of N_2 at the equator due to solar irradiance produced photoelectrons, as calculated by the 1D model (*top*) and glow model (*bottom*) (*Bailey et al.* [2002]).

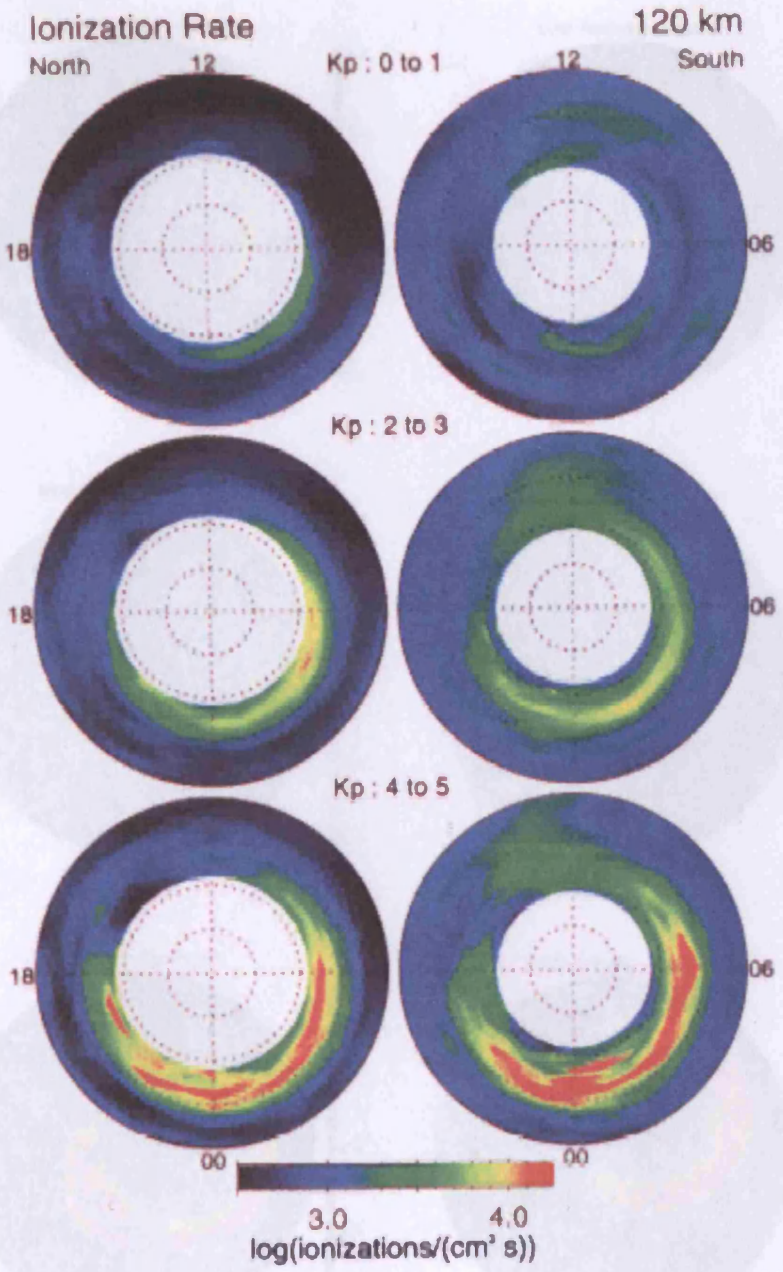


Figure 4.2 The total ionisation rate due to electron precipitation during low K_p (0-1), medium K_p (2-3), and high K_p (4-5) at an altitude of 120km computed using measurements from the PEM instrument on UARS. The centre of each plot is the magnetic pole, while the outer ring is at 50° latitude. *Ridley et al.* [1999].

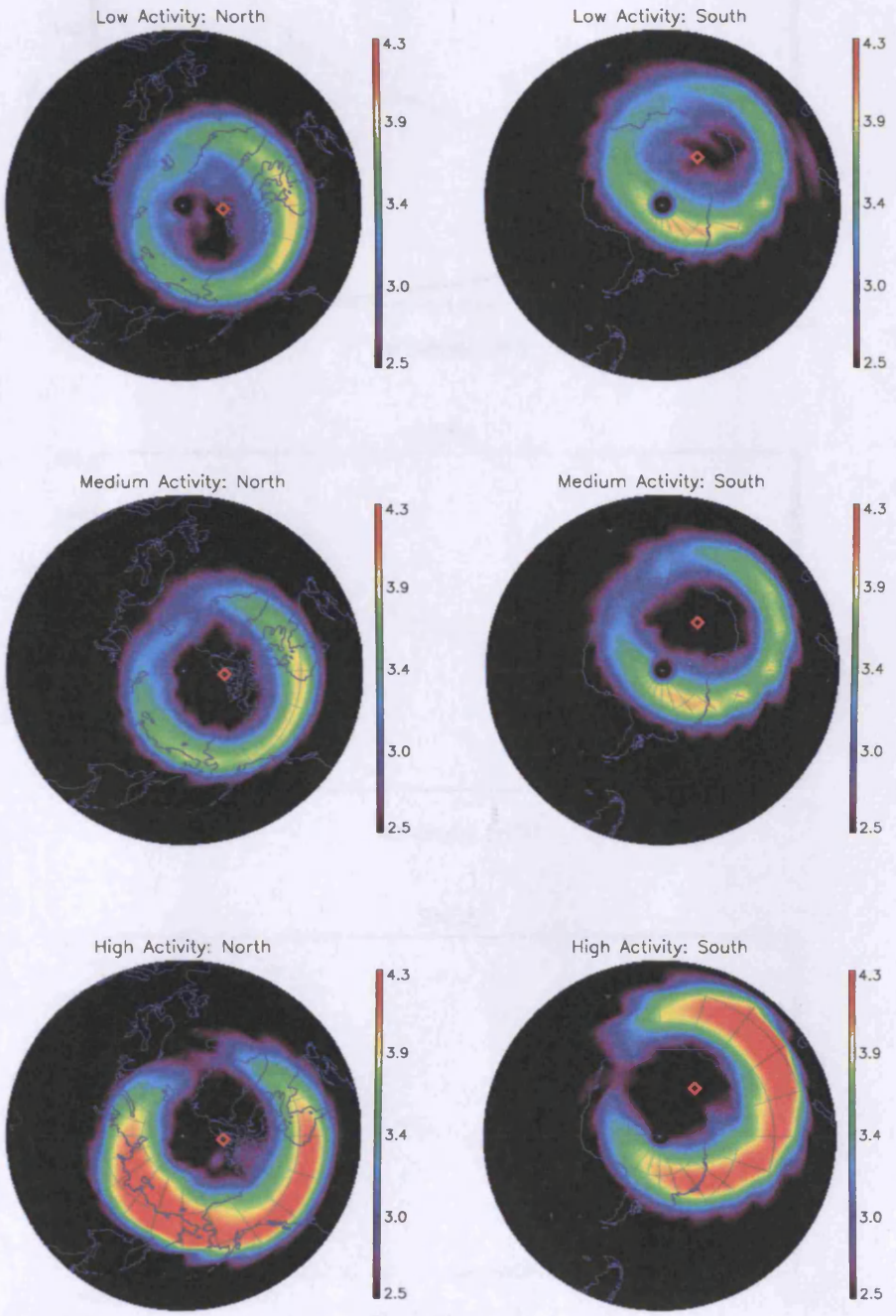


Figure 4.3 Total ionisation rate at 120km altitude in the northern and southern hemispheres during conditions of low (*top*), medium (*middle*) and high (*bottom*) geomagnetic activity as derived from the TIROS data. The centre of each plot is the geographic pole, while the outer ring is at 50° latitude. The red symbol shows the position of the geomagnetic pole in each hemisphere.

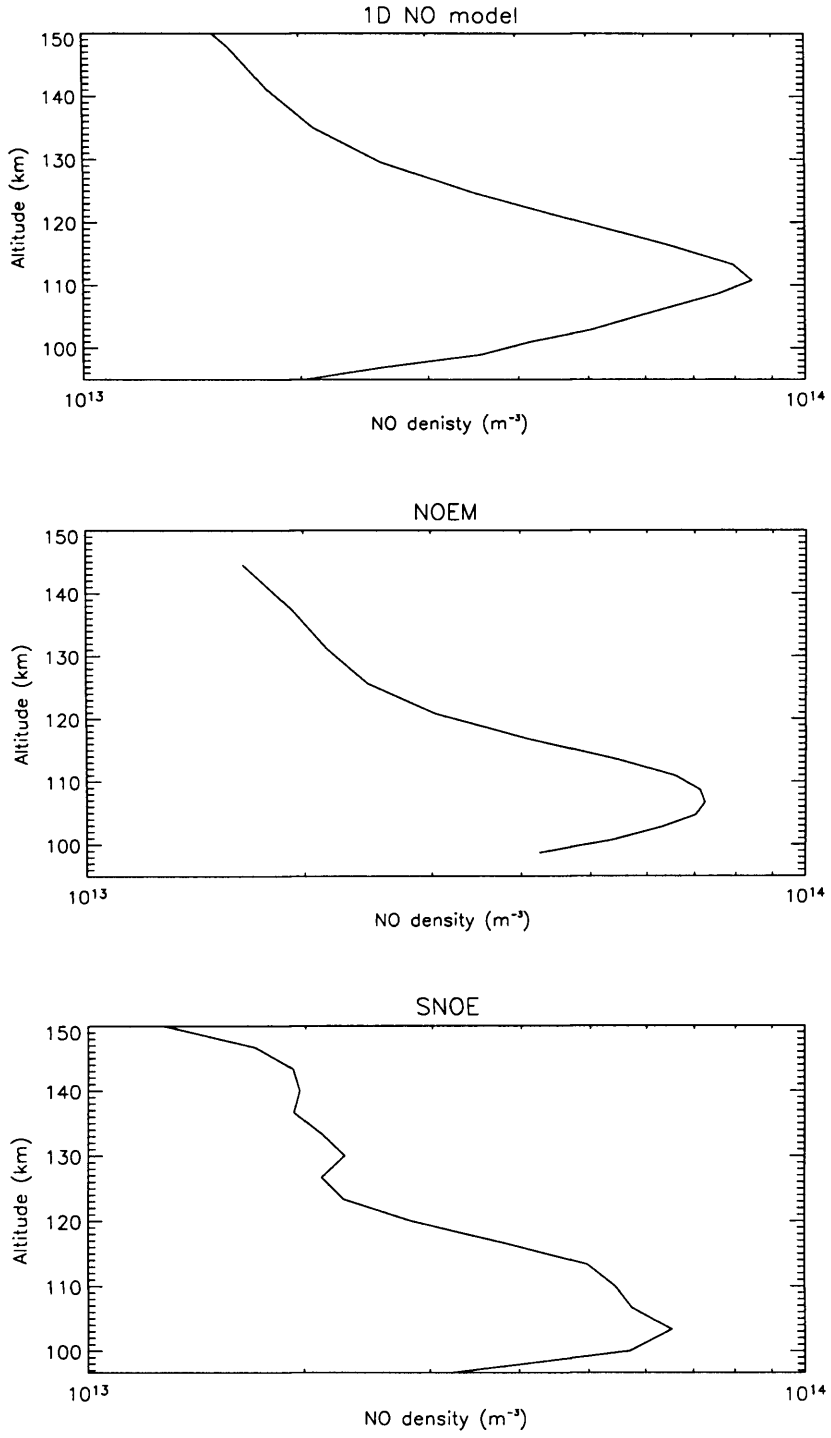


Figure 4.4 NO densities at the equator as calculated by the 1D model (*top*), NOEM (*middle*) and as given by SNOE (*bottom*) for conditions of low solar activity (F10.7=104), appropriate to equinox 1998.

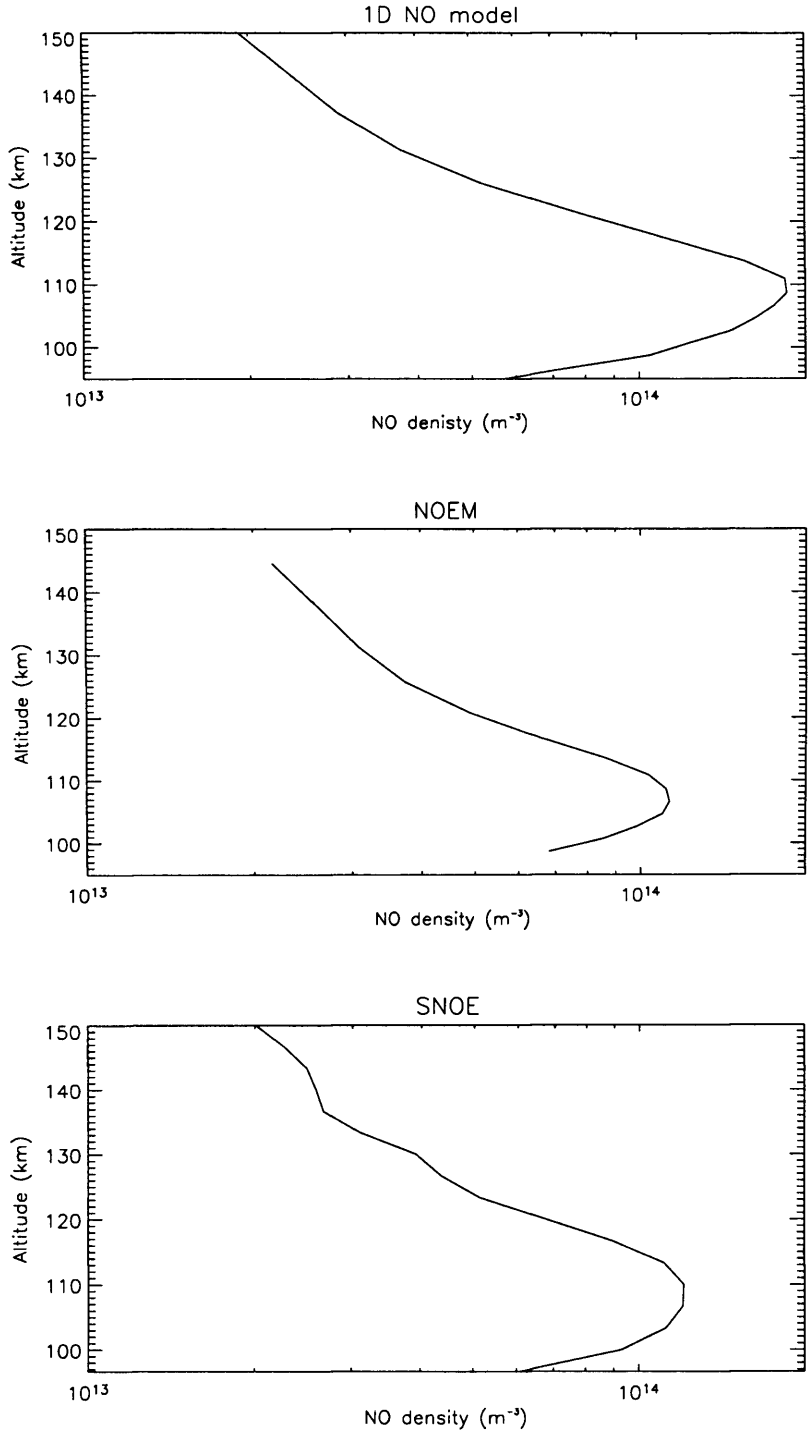


Figure 4.5 NO densities at the equator as calculated by the 1D model (*top*), NOEM (*middle*) and as given by SNOE (*bottom*) for conditions of high solar activity (F10.7=232), appropriate to equinox 2000.

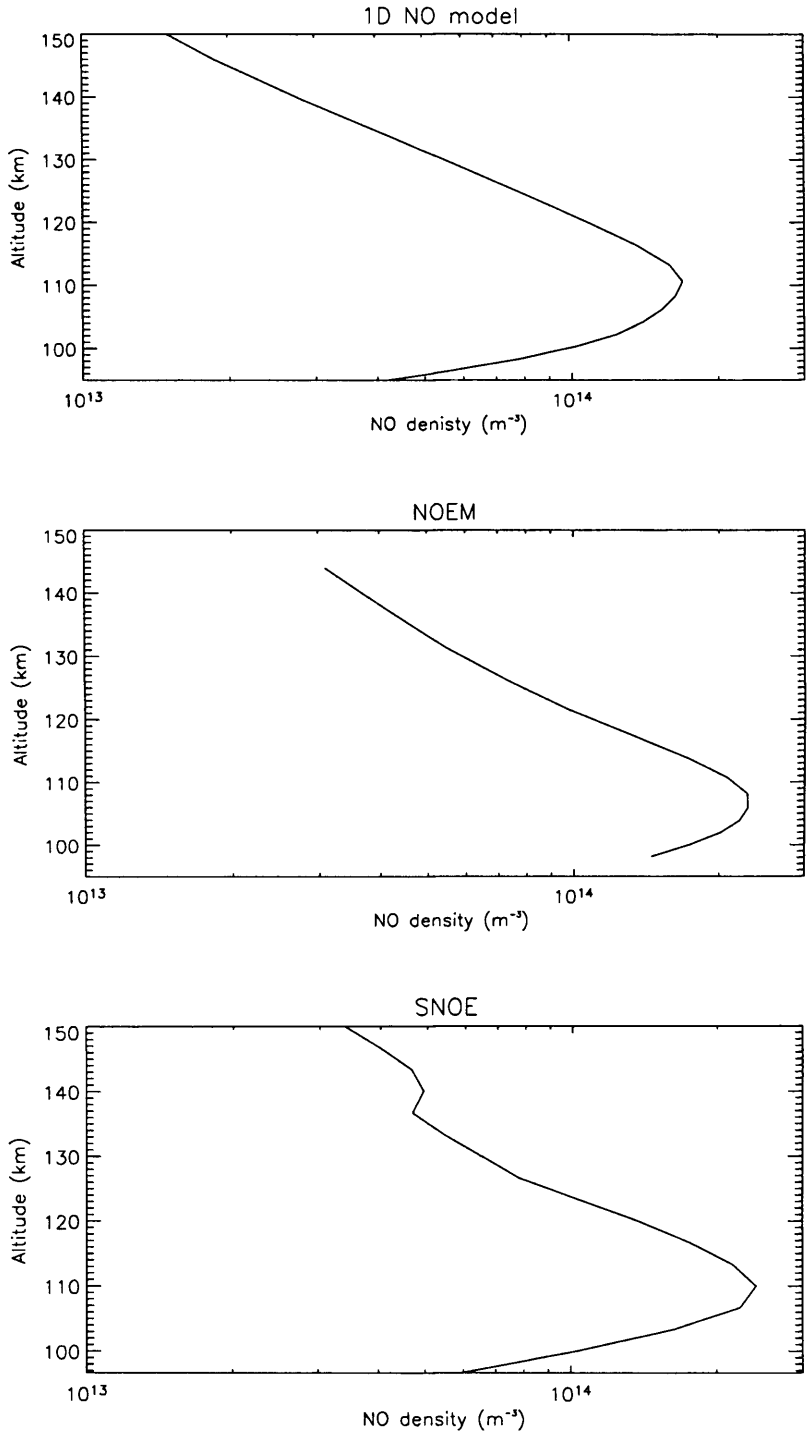


Figure 4.6 NO densities at 65° North as calculated by the 1D model (*top*), NOEM (*middle*) and as given by SNOE (*bottom*) for conditions of low geomagnetic activity (K_p = 1°), appropriate to equinox 1998.

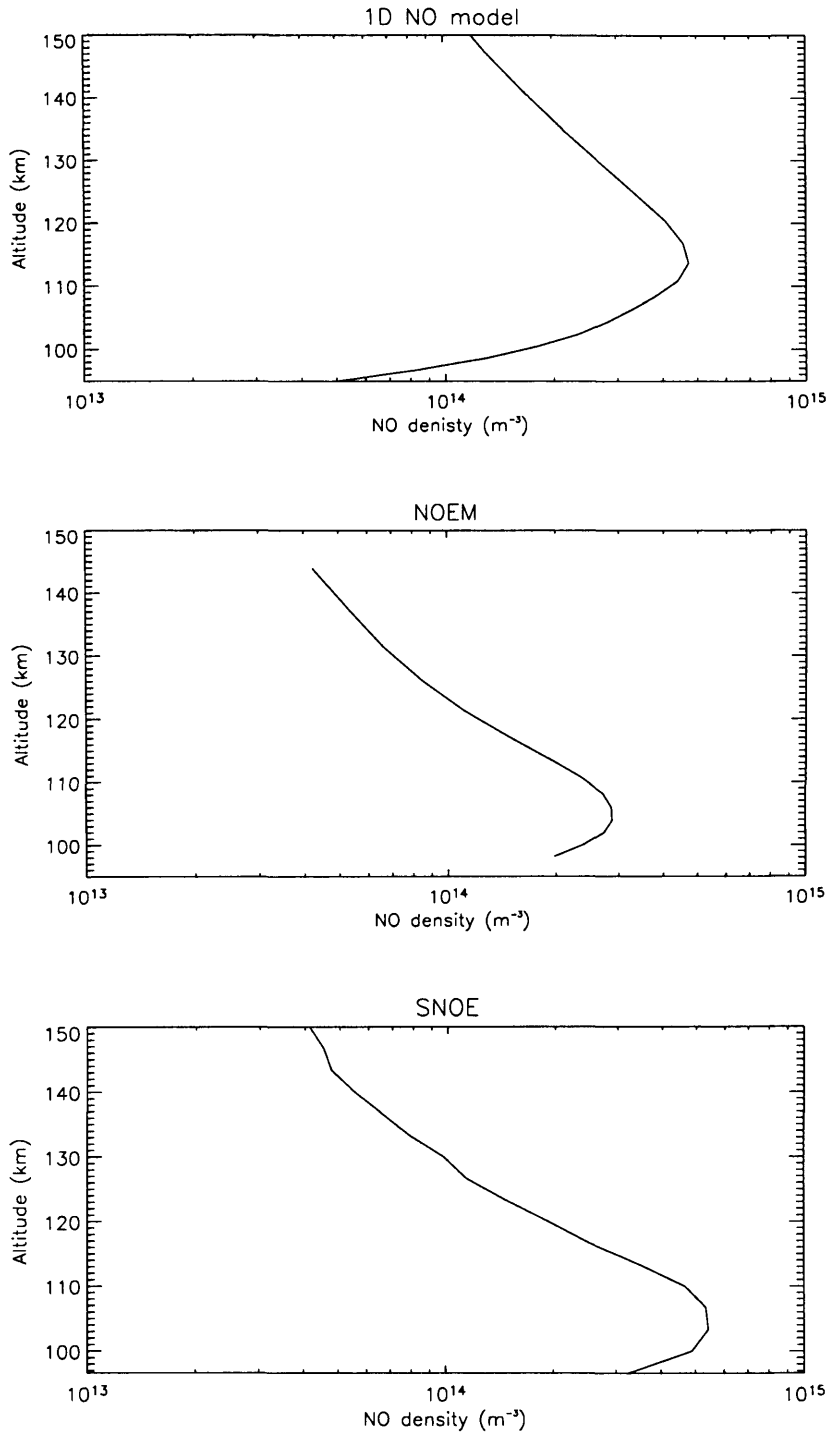


Figure 4.7 NO densities at 65° north as calculated by the 1D model (*top*), NOEM (*middle*) and as given by SNOE (*bottom*) for conditions of high geomagnetic activity ($K_p = 4^+$), appropriate to day 240 of 1998.

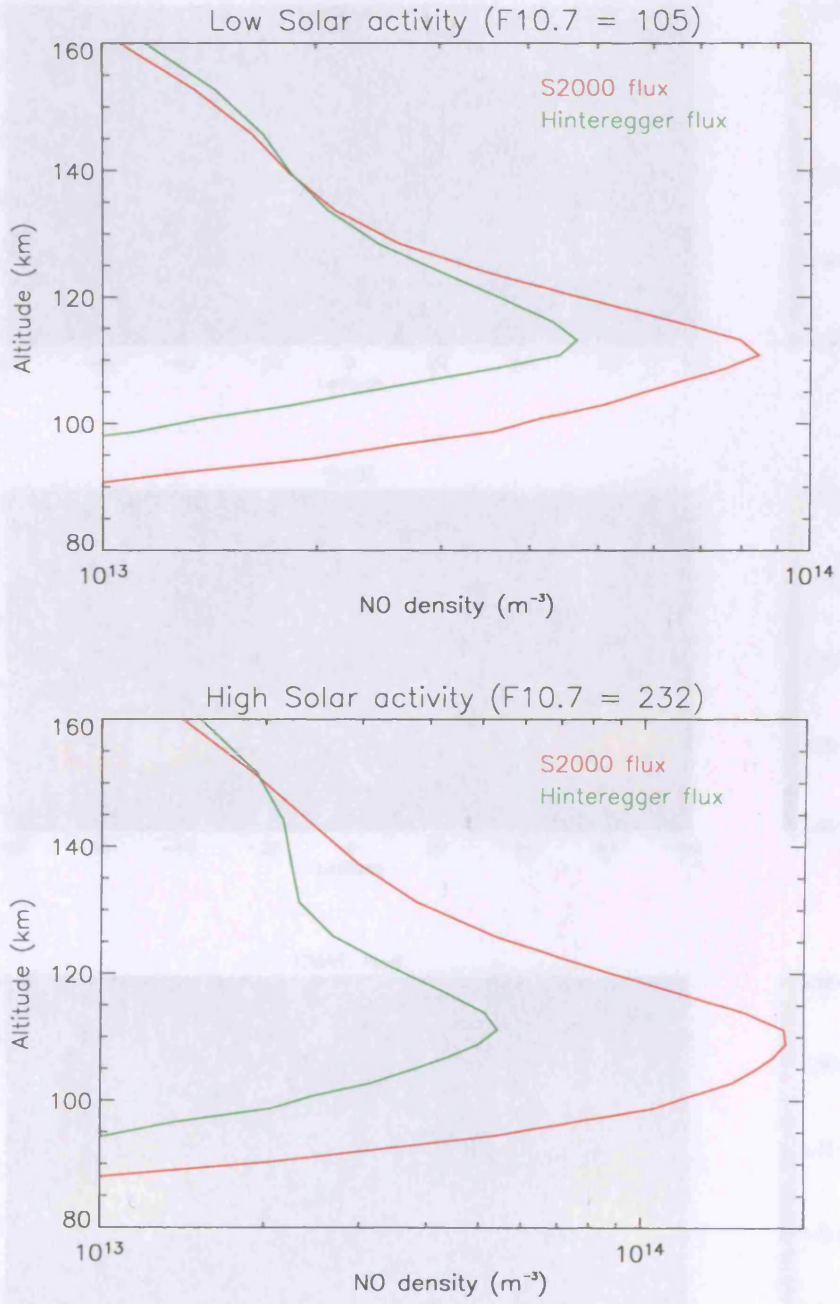


Figure 4.8 NO densities at the equator as calculated by the 1D model using solar flux data based on *Hinteregger et al.* [1981] for low (F107 = 105) and high (F107 = 232) solar activity, northern spring equinox.

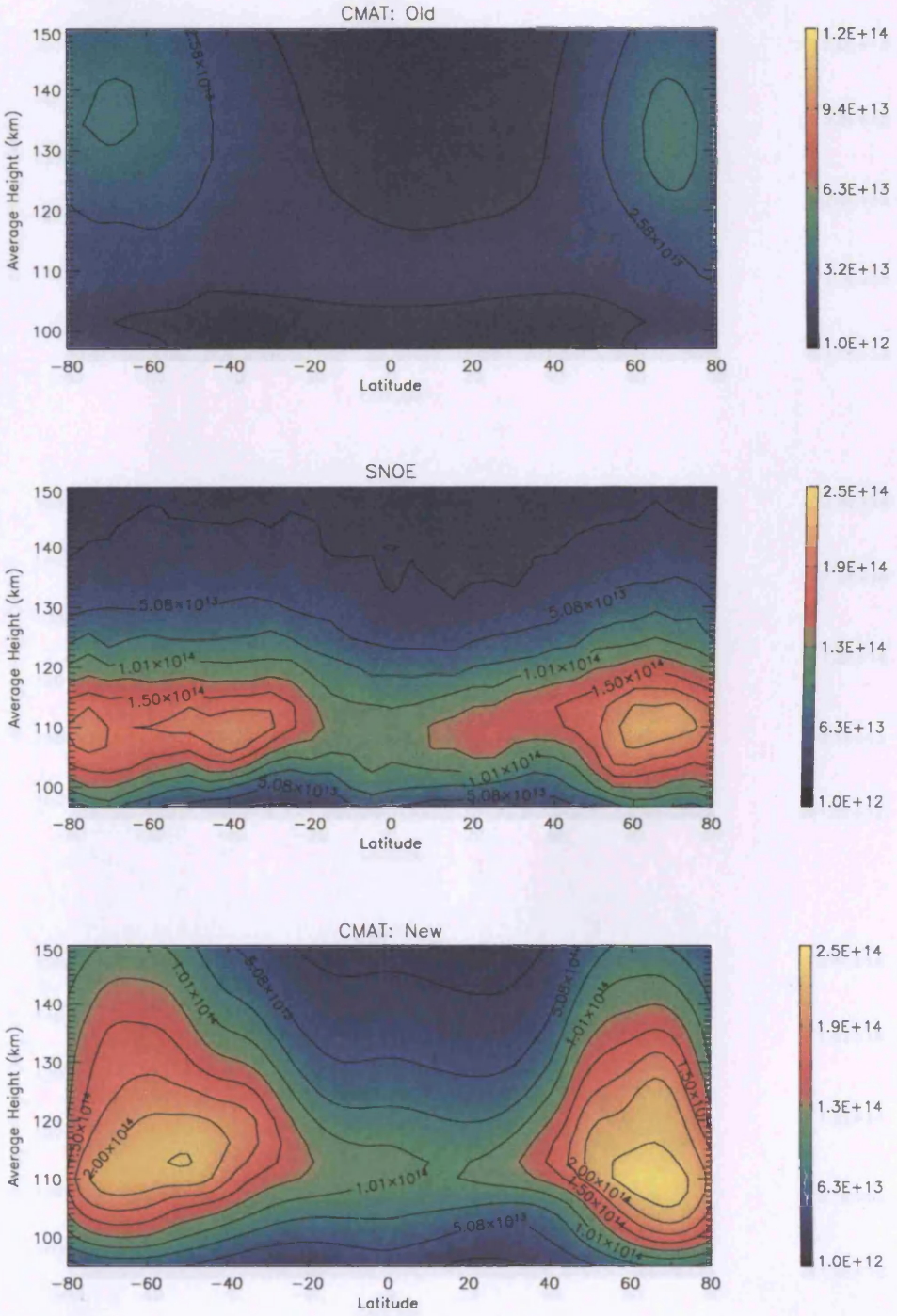


Figure 4.9 Zonal mean Nitric Oxide densities at 11am local time as given by CMAT before inclusion of the 1D NO model (*top*), SNOE (*middle*) and CMAT after inclusion of the 1D model (*bottom*). Conditions are as for day 266 of 1999, F10.7 = 137, $K_p = 4^o$. Note the factor of 2 difference in scale in the top plot.

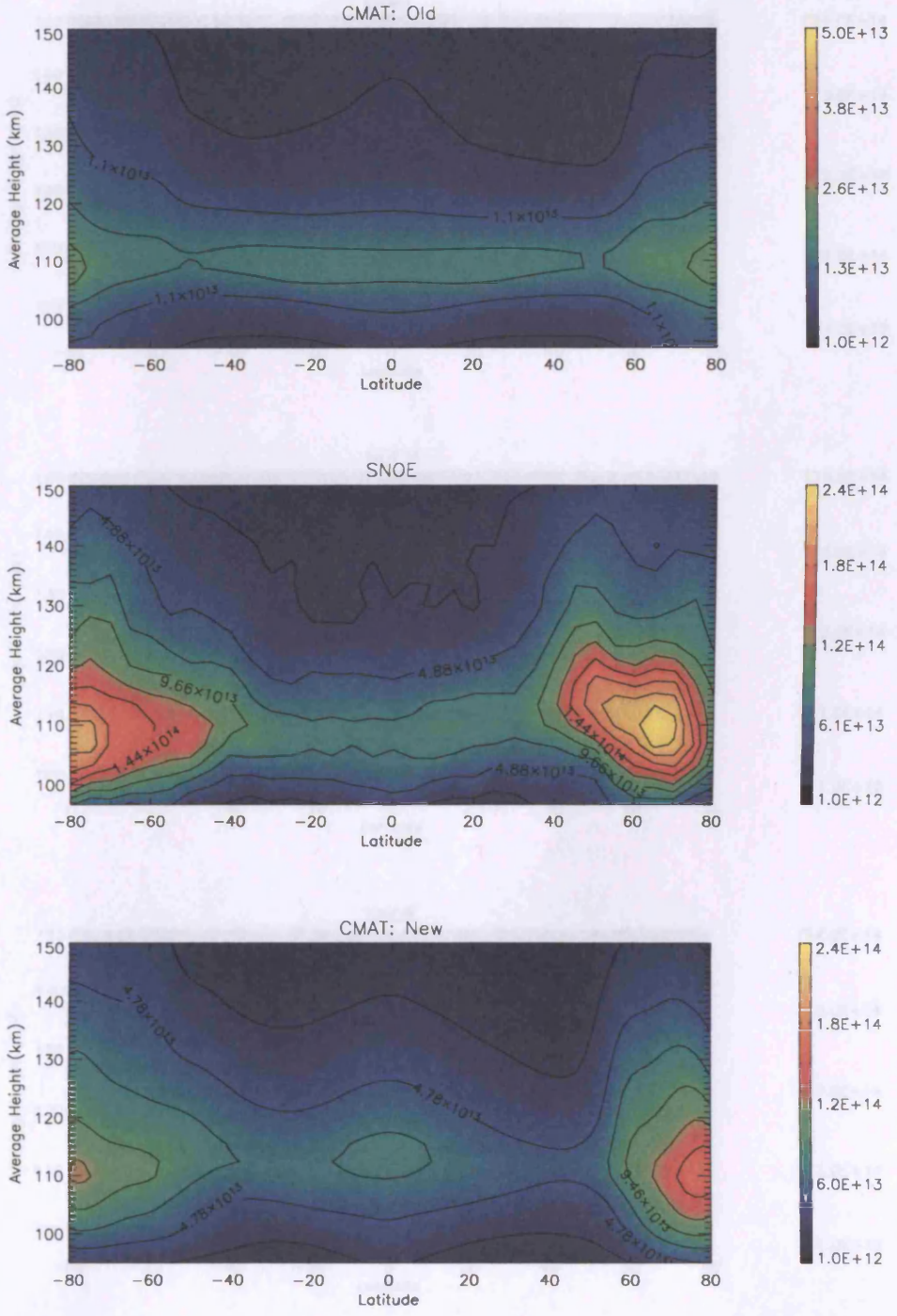


Figure 4.10 As for Figure 4.9 but geophysical conditions appropriate to day 79 of 1998. F10.7 = 126, K_p = 1°. Note the factor of 4 difference in the scale of the top plot.

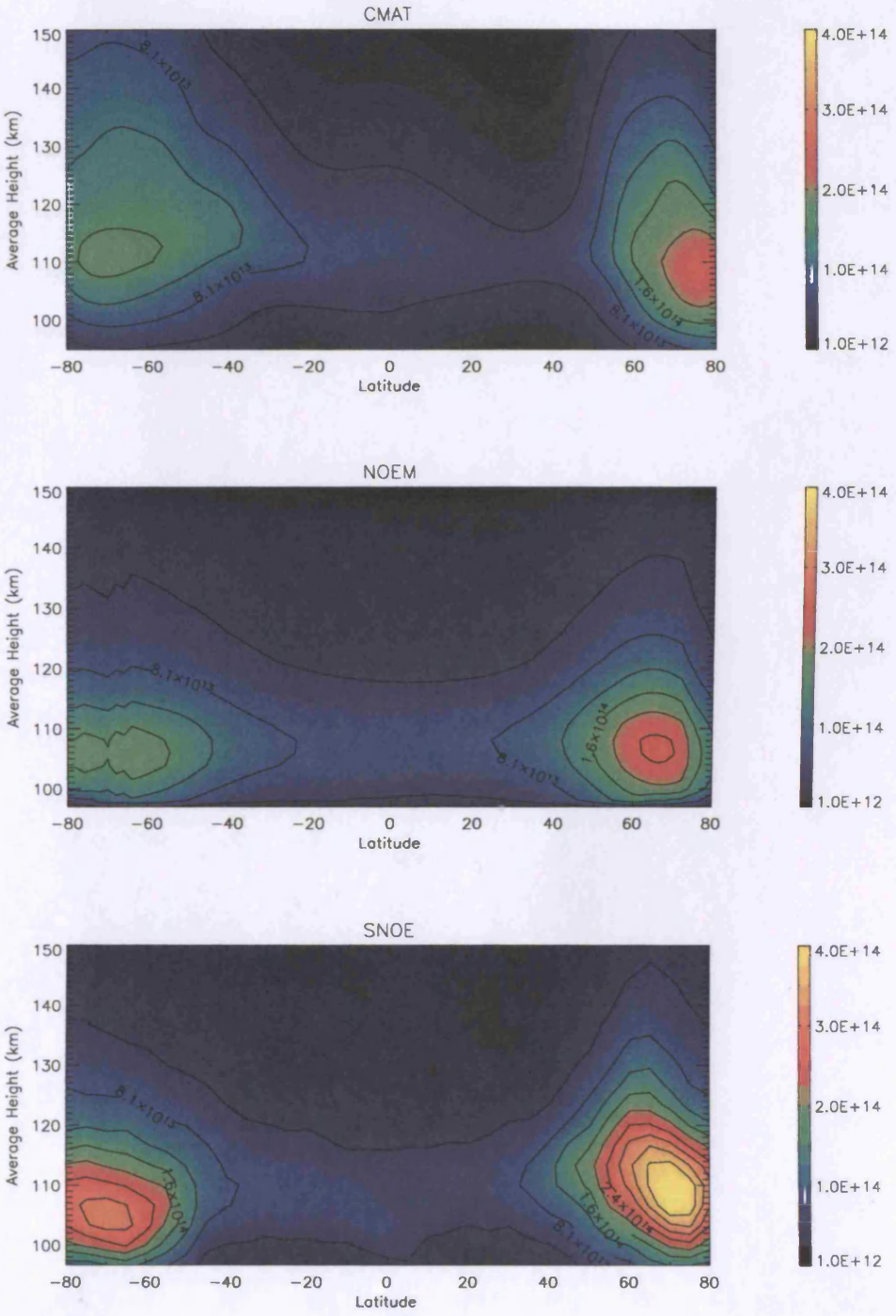


Figure 4.11 Zonal mean NO densities at 10:30am LT as calculated by CMAT (*top*), NOEM (*middle*) and as given by the SNOE satellite (*bottom*). Conditions are as for day 72 of 1998 when solar activity was low (F10.7 = 104) and geomagnetic activity was moderate to high ($K_p = 3^-$).

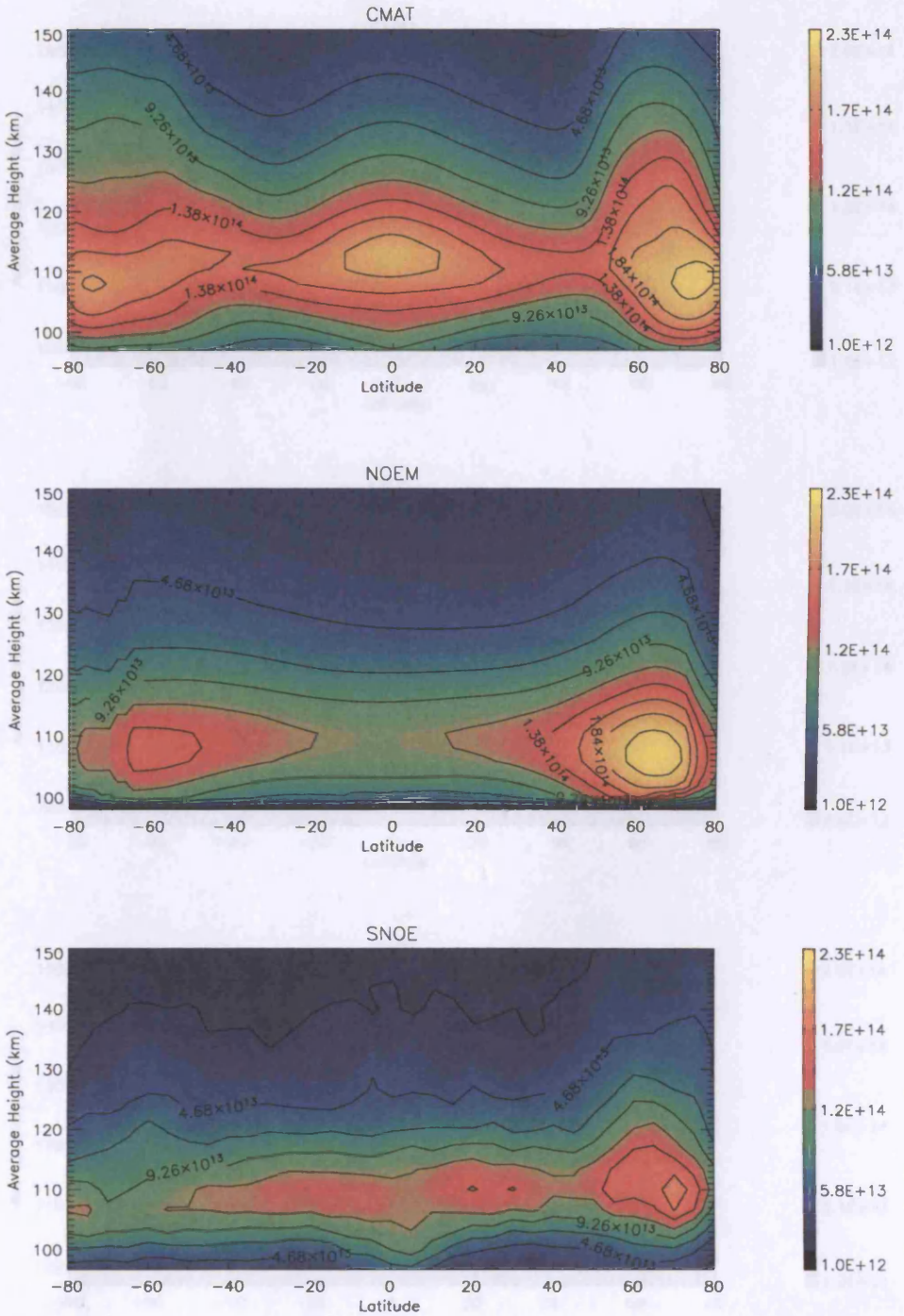


Figure 4.12 Zonal mean NO densities at 10:30am LT as calculated by CMAT (*top*), NOEM (*middle*) and as given by the SNOE satellite (*bottom*). Conditions are as for day 82 of 2000 when solar activity was high (F10.7 = 232) and geomagnetic activity was moderate to high ($K_p = 3^-$).

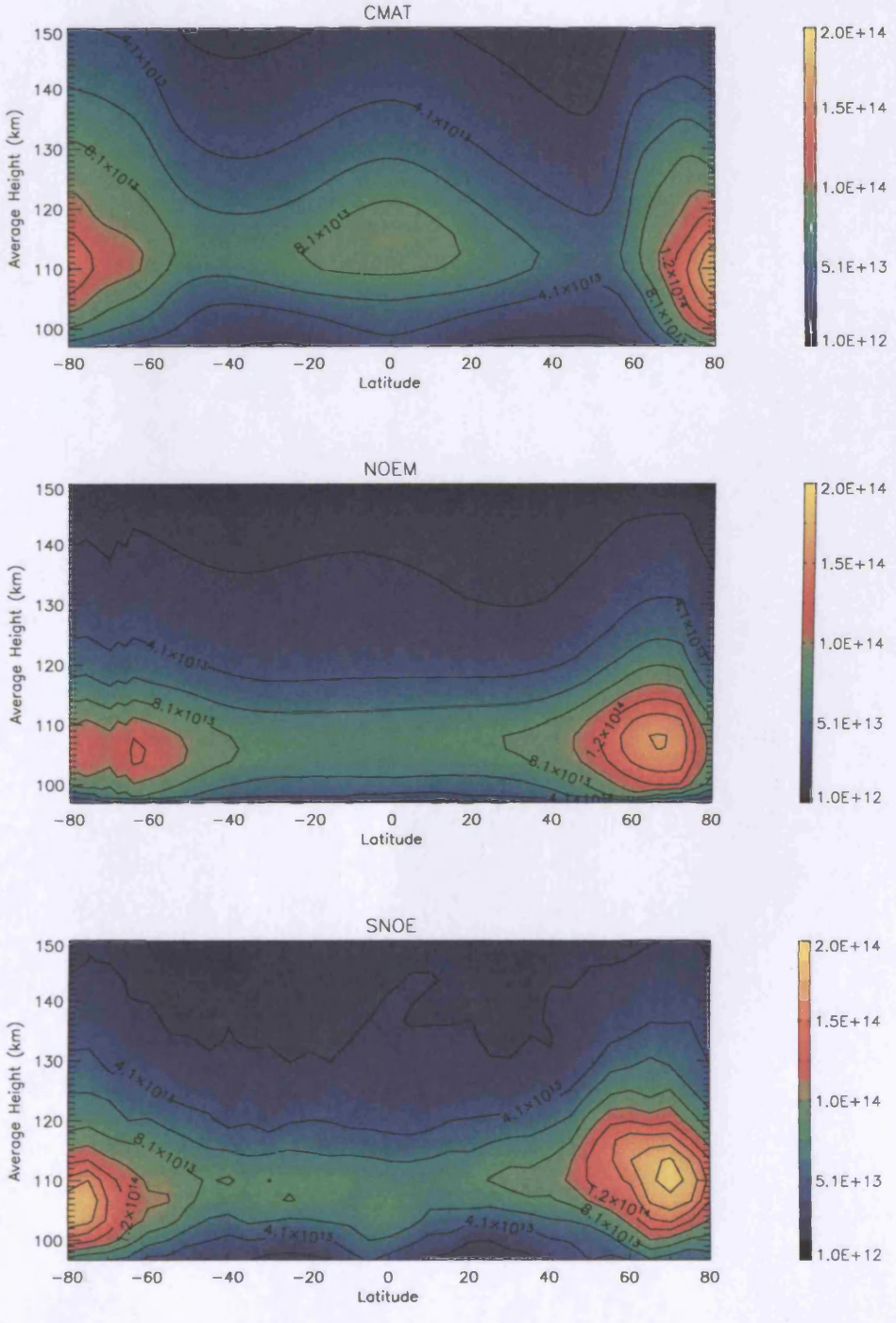


Figure 4.13 Zonal mean NO densities at 10:30am LT as calculated by CMAT (*top*), NOEM (*middle*) and as given by the SNOE satellite (*bottom*). Conditions are as for day 78 of 1998 when solar activity was moderate (F10.7 = 124) and geomagnetic activity was low (K_p = 1°).

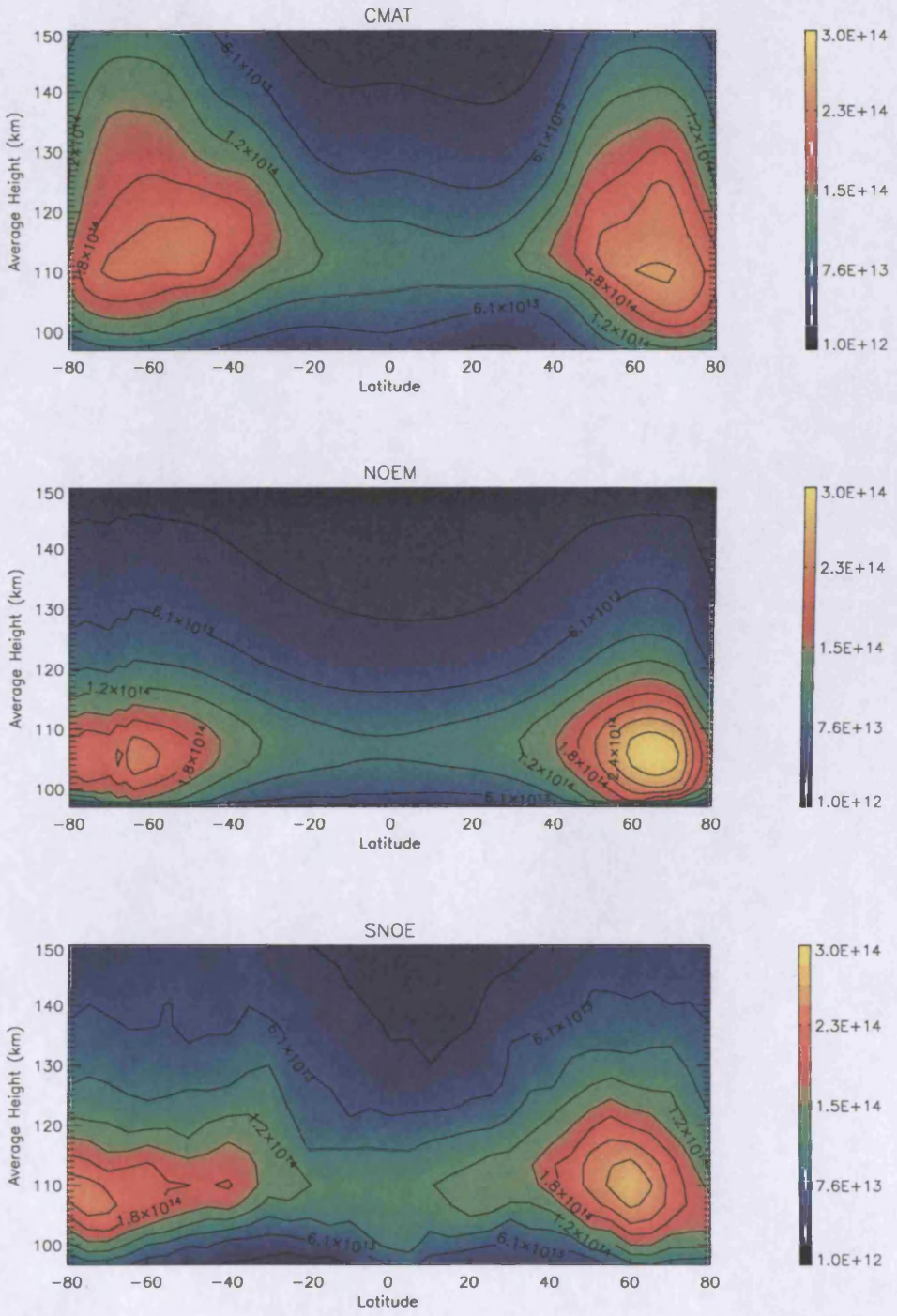


Figure 4.14 Zonal mean NO densities at 10:30am LT as calculated by CMAT (*top*), NOEM (*middle*) and as given by the SNOE satellite (*bottom*). Conditions are as for day 266 of 1999 when solar activity was moderate (F10.7 = 138) and geomagnetic activity was high ($K_p = 4^o$).

CHAPTER V. THE TEMPORAL AND SPATIAL RESPONSE OF NITRIC OXIDE AND THERMOSPHERIC TEMPERATURE TO HIGH GEOMAGNETIC ACTIVITY.

5.1 INTRODUCTION

A well established link between auroral energy input and Nitric Oxide concentration has been demonstrated by numerous observational and modelling studies (e.g. *Gerard and Barth* [1977], *Barth* [1992], *Solomon et al.* [1999], *Baker et al.* [2001], *Barth et al.* [2003]). An increase in the flux of precipitating electrons leads to an increase in ionisation that in turn leads to enhanced production of NO. During periods of high geomagnetic activity, Joule heating also contributes to increased NO production via the temperature sensitive reaction between ground state atomic nitrogen and molecular oxygen. Excess NO produced at auroral latitudes is thought to be transported by meridional winds, leading to increased concentrations of NO at mid and possibly even low latitudes. Considering the chemical and radiative properties of NO, the density distribution may give some indication of the temporal and spatial distribution of energy deposition in the thermosphere. Thus, we may gain some insight into the way auroral energy is distributed following periods of high geomagnetic activity.

In a recent study by *Barth and Bailey* [2004], data from the SNOE satellite was compared with NO densities calculated by a 1D photochemical model that did not include auroral energy sources. While a good match existed between modelled and observed NO abundances at low latitudes, the SNOE data revealed the presence of excess NO concentrations poleward of ±30°. The proposed source of this excess was aurorally produced NO that had been transported equatorward. *Marsh et al.* [2004] also proposed that aurorally driven variability in NO density exists poleward of ±30°. Both the 1D modelling study of *Barth and Bailey* [2004] and the observational study of *Barth et al.* [2003] suggested that at times of large geomagnetic storms, aurorally produced NO can be transported all the way to the equator. While there has been much speculation on the latitudinal extent of NO transport following periods of high geomagnetic activity, to date, no 3D modelling studies have been performed.

A key property of the thermosphere that is strongly influenced by NO abundance is the thermal response to periods of enhanced geomagnetic activity. 5.3μm radiative emission by excited NO is an important cooling mechanism in the lower thermosphere (*Maeda et al.* [1992]) and as such, enhanced levels of NO created by high auroral activity will act to increase cooling in the recovery phase of a storm. As well as increasing NO production, enhanced auroral energy deposition influences the collisional excitation of NO, which occurs primarily via inelastic collisions with atomic oxygen. A change in the abundance of O is consequently accompanied by a change in the amount of excited NO molecules. Increases in the background temperature associated with Joule heating can act to increase the rate of collisional excitation. Recent observations of infrared radiance by the SABER instrument onboard the Thermosphere-Ionosphere Mesosphere Energetics and Dynamics (TIMED) satellite show dramatic increases in 5.3μm emission during periods of enhanced solar and magnetic activity (*Mlynczak et al.* [2003]). This radiative emission is one way in which solar energy is converted in the atmosphere from one form to another.

From the processes described above, it follows that any attempt to simulate the thermal response of the atmosphere to auroral forcing must include a reasonable representation of NO chemistry and transport. *Maeda et al.* [1989] used a zonally averaged thermospheric model to show that the activity dependent variability of NO must be taken into account in order to reasonably simulate both the response and recovery of thermospheric temperatures to changing geomagnetic activity. When unrealistic NO densities were used, the modelled temperature response was drastically overestimated and the relaxation time following subsidence of activity, considerably slower than that of observations. Modelling studies of the ionospheric response to geomagnetic forcing also exhibit prolonged recovery times in simulated peak electron densities (*NmF2*) when odd nitrogen chemistry is not well represented (*Fuller-Rowell et al.* [2000]). *Wells et al.* [1997] showed that when the effects of NO cooling were included in the CTIP model, via application of an *effective heating efficiency*, the representation of quiet time neutral temperatures, composition and maximum plasma frequency (*foF2*) were substantially improved in the summer hemisphere.

While numerous modelling studies of the response of the thermosphere and ionosphere to geomagnetic storms have been performed (e.g., *Fuller-Rowell et al.* [1994], *Fuller-Rowell et al.* [1996c], *Fujiwara et al.* [1996], *Field et al.* [1998], *Burns et al.* [2004]), few have looked in detail at the thermal response and recovery timescales. The study of *Field et al.* did include a discussion of the variation in upper atmospheric neutral gas temperature following a geomagnetic storm. Advection of NO was however not included in the calculations. In this chapter, the thermal response of the thermosphere to enhanced geomagnetic activity is simulated using the CMAT model, which includes a fully self-consistent calculation of odd nitrogen chemistry and transport.

In the following sections, the results from three CMAT GCM experiments are described. In all simulations, the distribution of particle precipitation and energy flux are taken from the statistical model of *Fuller-Rowell and Evans* [1987]. The cross-polar cap electric field is taken from the model of *Foster et al.* [1986]. The first experiment illustrates the latitudinal range over which aurorally produced NO is distributed under conditions of moderate geomagnetic activity. In the second experiment, a simple '*quasi-step*' geomagnetic storm was generated. The resultant profiles of NO density are presented and compared to equivalent profiles produced under conditions of moderate activity. Simulations were performed with and without the effects of advection in order to assess the impact of transport processes on the distribution of aurorally produced NO. Thermospheric temperature variations arising from the storm forcing are also discussed. In the final simulation, an 11-day period from 23rd October to 3rd November 2003 is simulated using the CMAT and CTIP GCMs. During this period the Earth experienced one of the largest geomagnetic storms ever recorded. Neutral temperatures calculated by the models are compared with those measured by the Fabry-Perot Interferometer at Kiruna, northern Sweden. The meridional distribution of CMAT simulated temperature variations at 250km altitude are presented, along with the peak nitric oxide distribution resulting from the intense auroral forcing.

5.2 STUDY 1: MODERATELY ACTIVE CONDITIONS

Three CMAT model simulations have been performed. Each for a total of 40 model days such that the model is in a stable state where the diurnal variability of one day is a repeat of that from the previous day. The first simulation was a control run with no particle precipitation inputs. The second simulation was run at a moderate level of auroral activity with particle precipitation and electric field inputs defined by the statistical models of *Fuller-Rowell and Evans* [1987] and *Foster et al.* [1986] respectively. Both the particle precipitation and electric fields are specified according to a single power index (PI) which describes the total amount of particle energy delivered to an entire auroral hemisphere over half an orbit by the NOAA satellite (about 45 minutes), as described in *Foster et al.* [1986]. In this case, the power index has been set to 5, which is representative of a K_p of 2^+ and an estimated hemispheric power input of between 10 and 16 GW. In order to identify the effects of transport on aurorally produced NO, the third simulation has been run with auroral inputs appropriate to power index 5, but with horizontal and vertical advection terms set to zero in the minor species continuity equation.

All other geophysical conditions in the three simulations are the same and are appropriate to low solar activity (F10.7 = 105) during northern spring equinox (day 79 of the year).

5.2.1 LATITUDINAL EXTENT OF AURORALLY PRODUCED NO

Latitude versus height profiles of nitric oxide number density as calculated by the CMAT model for the three cases described above are shown in Figure 5.1. All plots are for a longitude of 0° at 12:00 UT (i.e. 12:00 LT) and demonstrate the typical morphology of NO number density resulting from each model simulation. Comparison of the density profiles created with and without auroral forcing (top and middle plots respectively) reveals enhanced levels of NO when auroral inputs are included. The high latitude peaks occur at 110 and 118km altitude in the northern and southern hemispheres respectively, the northern hemisphere peak being slightly greater in magnitude than that in the south. Differences in the locations and magnitude of these peaks arise from asymmetries in the auroral energy inputs in each hemisphere. In the

case where auroral forcing is included but advection is set to zero (bottom plot), the high latitude peaks are slightly greater in magnitude and more localised than when transport effects are included. This is to be expected as vertical and horizontal winds will act to carry NO produced in the auroral zones, away from the production site. NO also spreads further poleward from the equator into mid latitude locations when transport is included.

A more detailed appreciation of the equatorward spread of aurorally produced NO can be gained from Figure 5.2 which shows difference plots of NO number density as calculated in simulations 1 and 2 described above. The plots show the difference in calculated NO density on a constant pressure level, at approximately 110km altitude where the peak abundance is expected to occur. Profiles for four different universal times are shown, as differences in day-side and night-side magnetic reconnection results in different auroral forcing throughout the day. At 00:00 UT, enhanced NO concentrations are present in the high latitude regions of the northern and southern hemispheres when auroral forcing is included in CMAT. Significant amounts of aurorally produced NO are present at latitudes above approximately 50° in the north and below -30° in the south. A strong variation in the meridional and zonal extent of the enhanced NO region exists as a result of the offset between the geographic and geomagnetic poles. The aurorally produced NO reaches further toward the equator in the southern hemisphere where the offset between geographic and geomagnetic poles is largest. Spatial and temporal variations in the auroral oval also contribute to the horizontal and temporal variation in NO distribution. The largest difference in calculated NO occurs close to -80° in the southern hemisphere where an excess NO density of $2.5 \times 10^{14} m^{-3}$ is present. At 06:00 UT, the northern hemisphere high latitude NO is more varied and extends slightly further equatorward than at 00:00 UT. The peak in aurorally produced NO in the southern hemisphere is however slightly reduced, as is the spatial variation. At 12:00 UT the maximum NO enhancement occurs in the northern hemisphere where a peak of nearly $2.5 \times 10^{14} m^{-3}$ is present close to 80° latitude. A similar situation is seen at 18:00 UT although the southern hemisphere high latitude NO enhancement is greater than at mid-day. Over the whole

day, the latitudes furthest towards the equator to be affected by significant amounts of aurorally produced NO, under conditions of moderate geomagnetic activity, are approximately -30° in the south and 45° north.

5.2.2 INFLUENCE OF ADVECTION ON AURORALLY PRODUCED NO

In order to demonstrate the amount of influence advection has on the horizontal distribution of aurorally produced NO, plots of the difference in calculated NO density between simulations 2 and 3 are shown in Figure 5.3. Both of these model runs used auroral energy inputs appropriate to a K_p of 2^+. Simulation 2 included the effects of horizontal and vertical species advection whereas simulation 3 did not. The plots are derived from calculated NO number densities at a longitude of 0° at four different universal times. Positive values illustrate regions where NO densities are enhanced by transport, whereas negative values show areas of depleted NO. In each of the plots, NO concentrations are enhanced at mid-latitudes when advection terms are included in the species continuity equation. Depending on time of day, southern hemisphere NO densities at the peak altitude of approximately 110km, are reduced at latitudes poleward of about –60° to -70°. At all UTs shown, NO densities are enhanced between –50° and –60°, extending to -30° at 06:00 UT. In the northern hemisphere a similar situation occurs although the region of enhanced NO is further poleward. Aurorally produced NO is being transported away from the region of production, to latitudes equatorward of the auroral oval. The vertical, horizontal and temporal extent of the NO production and transport will be strongly dependent on the position and structure of the auroral oval. Variation in the distribution and magnitude of high latitude energy inputs will influence the production and transport of species in and from that region. Joule heating and associated winds will depend on the nature and location of auroral energy inputs and as such so will the distribution of NO. The profiles shown here will therefore be highly variable with auroral energy inputs. A general appreciation of the extent of NO transport and horizontal distribution under conditions of moderate geomagnetic activity can however be gained from this study.

At low latitudes, the NO number density may be enhanced or diminished depending on the altitude and time of day. The cellular structure is indicative of tidal influence, as is

the vertical distance between peaks and troughs, which matches the wavelength of the diurnal tide (approximately 25 to 30km). The perturbations are strongest at the equator where tidal variations in the vertical wind tend to maximise. The influence of vertical tidal motions on the distribution of NO has been investigated by *Marsh and Roble* [2002] who confirmed that sunrise/sunset asymmetries in satellite measurements of NO are predominantly due to tidal forcing.

5.3 STUDY 2: MODEL RESPONSE TO HIGH AURORAL ACTIVITY

The aim of this study was to assess the changes in modelled nitric oxide density in response to a period of high auroral energy influx. The timescale of post storm thermospheric temperature recovery, which is strongly influenced by NO concentration, has also been assessed. As in study 1 described above, the CMAT model was first run to a steady state condition over a period of 40 days using conditions appropriate to northern spring equinox (day 79). An F10.7 of 105 was used, along with statistical particle precipitation and high latitude electric fields patterns associated with a power index (PI) of 5 (Kp = 2^+). This represents conditions of moderate auroral activity. From this starting point, a 6 day run was performed that used auroral energy inputs appropriate to a PI of 5 for the first 24 hours, after which the PI was increased linearly from 5 to 10, over a period of 1 hour. Activity level 10 is representative of high auroral activity. Whilst using this PI, the hemispheric power input was set to 125GW after *Fuller-Rowell et al.* [1994, 1996c], resulting in auroral energy inputs equivalent to a K_p of approximately 6^-. Auroral inputs remained at this level for 12 hours, after which the PI was relaxed back to level 5. This level of activity was then held constant at a PI of 5 for the remainder of the 6-day period. The table below summarises the time history of auroral power inputs to the model as a function of hours from the start of the 6-day run. Note that CMAT days start at 12:00 UT so hour 0 to 1of the run occurs from 12:00 to 13:00 UT and so on.

Hour	*PI*	K_p
0 to 24	5	2^+
24 to 25	5 to10	2^+ to 6^-
25 to 37	10	6^-
37 to 38	10 to 5	6^- to 2^+
38 to 120	5	2^+

Table 5.1 Time history of auroral power inputs to the model as a function of hours from the start of the 6-day simulation.

From the same starting point, a control simulation was also performed using auroral energy inputs characteristic of PI 5 throughout the full 6-day period.

The quasi-step variation of auroral power input utilised here is considerably simpler than the complex variations that occur in reality. Despite the apparent simplicity, this approach has been used in numerous modelling studies (e.g. *Fuller-Rowell et al.* [1994], *Fuller-Rowell et al.* [1996c], *Field et al.* [1998], *Fujiwara et al.* [1996], *Burns et al.* [2004]) and is sufficient to simulate the first order characteristics of a storm response.

5.3.1 RESPONSE OF NITRIC OXIDE TO STORM FORCING

Calculated nitric oxide number densities at 110km as a function of time and geographic latitude are shown in Figure 5.4. The time scale refers to hours from the start of the 6 day model simulation, hour 0 representing 12:00 UT on the first day. Profiles are shown for 0°, 90°, 180° and 270° geographic longitude. High energy particle precipitation occurs in the auroral oval, between approximately 65° and 75°

geomagnetic latitude (*Newell et al.* [1996]). Clearly seen are high latitude enhancements that increase and decrease with time as the high latitude geographic grid points move through the auroral oval. The largest nitric oxide densities are seen between -40° and -50° latitude in the southern hemisphere at a longitude of 180° where concentrations of approximately 3.5 x 10^{14} m^{-3} are visible. The maximum appears 2 days after the end of the storm period (24 – 38 hours). In the northern hemisphere, maximum densities of about 3.0 x 10^{14} m^{-3} occur between 50° and 60° latitude at a longitude of 270°. This maximum occurs 38 hours after the storm forcing has ended. The amount of NO present at any point on the globe is clearly highly dependent on the geographic location in relation to the region of increased auroral energy input.

Figure 5.5 shows difference plots of calculated NO density from the simulation including 12 hours storm forcing, and the control simulation that used constant medium activity forcing throughout the 6-day period. During the first 24 hours, auroral forcing in both runs was appropriate to a PI of 5. As such there is no difference between the NO densities calculated in each simulation for the first day. Throughout the storm forcing period that follows, depletion in NO concentration at high latitudes is visible in both hemispheres. Mid latitudes however undergo an increase in NO density as the auroral oval expands to lower latitudes. This is most visible in the southern hemisphere where the geomagnetic pole is located further equatorward. Excess NO created by the sudden influx of auroral energy starts to be transported both equatorward and over the poles by meridional winds, generated as a result of steep pressure gradients created by high latitude Joule heating. At 110km altitude, modelled meridional winds at high latitudes reach ±75ms^{-1} in response to Joule heating (not shown). At altitudes above 130km, these winds can reach over 300ms^{-1} and act to rapidly transport NO to mid and low latitudes. A pole to equator circulation is set up with upwelling at the high latitudes and downwelling at low latitudes (*Maeda et al.* [1989]). By the end of the storm forcing period (38 hours into the simulation), increased concentrations of NO are visible within 30° of the equator at 180° longitude. Over the following 36 to 48 hours, the region affected by aurorally produced NO extends to latitudes even closer to the equator.

The timescales and extent of horizontal species transport differ significantly with longitude. At 0° longitude, transport of aurorally produced NO is limited to latitudes poleward of approximately 30° in both hemispheres. The maximum enhancements appear around 30 hours after the storm forcing has ended. At 180° longitude, significant amounts of excess NO are present within 12° and 18° of the equator in the southern and northern hemispheres respectively. These maxima occur up to 40 hours after the end of the storm forcing period. In general, the mid latitude NO enhancements start to die away between 24 and 48 hours after storm forcing ends. At certain locations however, significant quantities of excess NO are present throughout the remainder of the 6 day period. Southern hemisphere enhancements of nearly 1.0×10^{14} m^{-3} are present at approximately -40° latitude, 180° longitude, 90 hours after relaxation of the storm.

One factor that is likely to have an impact on the timescale and magnitude of changes in high latitude NO density following high energy auroral inputs, is the local time at which the auroral forcing occurs. If particle precipitation leads to the production of excess NO in the sunlit portion of the atmosphere, photodissociation by solar UV will act to destroy that NO. *Barth et al.* [2001] suggested the lifetime of an NO molecule to chemical destruction under illuminated conditions is 19 hours. This will be the case at 0° longitude in the simulation performed here. Conversely, if excess NO is created in the night-time, as is the case for 180° longitude here, that excess NO will be longer lived and available for transport to lower latitudes.

The highly variable nature of atmospheric energy sources makes it extremely difficult to assess the lifetime of NO following production in, and transport from the auroral zones. The response of the atmosphere to constantly changing auroral and solar energy inputs is manifest in satellite observations of NO density. Even so, it is possible to use satellite data to gain some qualitative idea of the temporal and spatial extent of NO transport following periods of high geomagnetic activity. *Marsh et al.* [2004] noted a 1 day lag in the maximum correlation between changes in observed high latitude NO density and the K_p index of geomagnetic activity. A similar delay was noted by *Barth and Bailey* [2004], in relation to the transport of aurorally produced NO to the equator

following periods of intense geomagnetic activity. Both of these studies used zonal averages of data from the SNOE satellite, which comprises NO measurements from a single local time (approximately 10:30LT). While the results presented here confirm that there is a lag in the maximum response of NO density to high energy auroral forcing, they also show that the lag is dependent on geographic location in relation to the auroral oval, and possibly the local time at which auroral forcing occurs.

By zonally averaging the CMAT NO data presented in this study and isolating concentrations from a single local time of 11:00AM, the CMAT data can be presented in a format comparable with the zonally averaged SNOE dataset. When this is done, the maximum NO number densities are seen at high latitudes on day 3 of the simulation (not shown). The average PI on day 2 is 7, corresponding to a K_p of 3^+, whereas the average PI on day 3 of the simulation is 5, corresponding to a K_p of 2^+. It is therefore reasonable to conclude that when zonally averaged NO data is considered, at a single local time, the maximum response to changes in the K_p geomagnetic index occurs with a 1 day lag. At present, observational datasets of NO concentration do not have global coverage and it is only through using a 3D GCM such as CMAT that detailed studies of the spatial and temporal distribution of aurorally produced NO can be performed.

Following periods of high geomagnetic activity, *Barth and Bailey* [2004] observed that NO enhancements could remain at low latitudes for several days after the storm period. The CMAT simulations performed here suggest that at equinox, auroral energy can have an effect on high and mid latitude NO concentrations for several days after relaxation of a storm. The lifetime of excess NO concentrations will be highly dependent on the number of hours in which the enhanced region is exposed to sunlight. Thus season and latitude are key factors in determining the longevity of aurorally produced NO, along with any additional sources of NO that arise via ongoing transport processes.

5.3.2 THERMAL RESPONSE

Neutral temperatures as calculated by CMAT in the '*storm*' simulation that included high geomagnetic activity forcing between hours 24 and 38 of the 6 day period are presented in Figure 5.6. Temperatures are shown as a function of latitude and time, for altitudes of approximately 130 and 300km, at 0° and 180° geographic longitude. The time series show the changes in temperature from hour 0 to 144 of the simulation. Clearly visible at both altitudes is an immediate increase in neutral temperature at the onset of the high auroral activity period. The maximum temperature enhancements occur at different times depending on longitude and latitude but generally the maxima appear 8 to 12 hours after storm onset. At 130km altitude in the northern hemisphere, high latitude temperatures reach over 800K at 180° longitude. At 0° longitude the high latitude temperature maxima are considerably less and reach just over 650K in two limited locations. At 300km altitude, high latitude temperature maxima at 180° are again higher than those at 0° longitude, reaching approximately 1300K and 1200K respectively. While the maximum changes in temperature are seen in the high latitude auroral regions, equatorial enhancements are also seen approximately 6 hours after storm onset. At 130km altitude, equatorial temperatures of over 550K are present at both 0° and 180° longitude, those at 180° reaching nearly 580K. Equatorial temperatures at 300km altitude, 0° longitude, exceed 1070K approximately 6 hours after commencement of the storm period. The equatorial increase occurs later at 180° where temperatures reach just over 1075K 16 hours after storm onset.

The rapid transport of energy from high to equatorial latitudes may be related to large scale atmospheric gravity waves (AGWs), generated as a result of the high latitude energy injection. *Fujiwara et al.* [1996] suggested that these AGWs can transfer energy from high to low latitudes at speeds close to the speed of sound (for example roughly 670ms^{-1} at around 260km), predominantly by adiabatic compressional heating and/or expansive cooling processes. For the simulation presented here, both adiabatic and advective processes are expected to play a part in the meridional transport of energy.

A more detailed appreciation of the magnitude and temporal development of temperature enhancements at both 130 and 300km altitude can be gained by subtracting the temperatures calculated in the control simulation from those of the storm simulation, as shown in Figure 5.7. At 130km altitude, maximum temperature increases of 150K occur at 0° longitude and are centred on -70° geographic latitude in the southern hemisphere. At 180° the temperature increases are much greater, reaching over 250K at 70° latitude in the northern hemisphere. Low latitude increases of about 50K are present at both longitudes, occurring 6 and 16 hours after storm onset at 0° and 180° respectively. At high latitudes, significant temperature enhancements remain for up to 26 hours after relaxation of the storm at both longitudes shown. At low latitudes, recovery to quiet time temperatures is slightly faster, occurring around 18 hours after relaxation of the storm. At 300km altitude, temperature enhancements of just less than 370K are seen at both longitudes, 8 to 12 hours after storm onset. The maximum increases in low latitude temperatures occur approximately 12 hours after commencement of the storm at both 0° and 180° longitude and reach between 60K and 80K. The timescales of recovery to temperatures appropriate to moderate geomagnetic activity are slightly longer than those at 130km altitude, significant high latitude enhancements remaining for up to 30 hours at both 0° and 180° longitude. Low latitude temperature recovery takes approximately 16 to 20 hours, in line with the timescales seen at 130km.

It is clear from the data presented here that both and magnitude and temporal scales of temperature changes resulting from geomagnetic storm forcing are highly variable with geographic location. At high latitudes, temperature enhancements are greater in magnitude and longer lived than at low latitudes. These enhancements appear to be slightly longer lived at higher altitudes, where downward heat convection is the dominant cooling process. At both 130 and 300km altitude, high latitude temperatures respond rapidly to increases in auroral energy inputs, resulting in sharp gradients in temperature at storm onset and relaxation.

In order to further investigate the timescales of temperature relaxation at high latitudes following high geomagnetic activity, zonal mean values of temperatures calculated in

the control simulation have been subtracted from zonal mean temperatures calculated with high activity auroral forcing. This removes the effect of longitudinal and diurnal variations, meaning that storm driven changes in temperature can be easily identified. Figure 5.8 shows the difference in zonal mean temperatures calculated in each run at a latitude of 65° north, for altitudes of approximately 120, 180, 240 and 300km. At each altitude the response to increased auroral energy input is immediate, with sharp increases in temperature seen over the first four hours. At 180 and 240km altitude, this initial increase is followed by a small decrease, probably related to expansion of the auroral oval. Maximum temperature enhancements occur approximately 12 and 8 hours after storm onset at the two lower altitudes, and two high altitudes respectively. After this point, a sharp decrease in the temperature occurs at all altitudes, although the decrease at 180km is short lived and followed by yet another increase. At the end of the storm forcing period, an overall decrease in temperature occurs at all altitudes. The small increase in temperature present at 120km is again attributed to movement of the auroral oval in relation to geographic location. Perhaps the most interesting feature of the time series illustrated here is the difference between the temperature relaxation at 120km and that at higher altitudes. A relatively smooth relaxation to pre-storm conditions is seen at 120km, total recovery occurring approximately 60 hours after the end of the storm. In contrast, while the higher altitude temperatures reach their pre-storm values by hour 72 of the simulation (36 hours after the end of the storm), they are not fully stabilised and a small oscillation exists to the end of the simulation .

The impact of NO cooling on the timescales of temperature response to increased geomagnetic activity can be seen by comparing the CMAT response to that of a model that does not include NO cooling, such as the CTIP GCM. Figure 5.9 shows the difference between zonally averaged neutral temperatures at 65° north calculated by the CTIP model in a storm and control simulation, using identical conditions to those used for the CMAT model runs. While the rapid increase in temperature, in response to increased geomagnetic activity, is similar to that seen in the CMAT simulation, the maximum temperature enhancement predicted by the CTIP model at 65° latitude is considerably greater. This is most pronounced at 180km where the zonally averaged

temperature enhancements at 65° latitude calculated by CMAT and CTIP are approximately 100K and 200K respectively. There are many differences in the way atmospheric dynamics and energetics are calculated in these two models. Consequently it is hard to pinpoint the cause of the temperature differences. One possible explanation lies in the difference in vertical resolution between the two models, CTIP having a vertical grid spacing of 1 scale height whereas CMAT uses 1/3 scale height. This will affect the altitude of energy deposition. The different solar spectra used in the two models will also lead to large differences in calculated thermospheric temperatures. It is worth noting that the CTIP model does not include a mechanism for NO radiative cooling but this is unlikely to have a dramatic effect on the initial temperature increase that occurs over the first few hours of high auroral activity.

The relaxation time to pre-storm temperatures is dramatically increased in the CTIP simulation. At 120km, the daily zonal mean temperature at 65° north is still enhanced by over 20K at the end of the six day simulation, 106 hours after the end of the storm forcing period. At higher altitudes, enhancements of around 40K remain at the end of the 6 day period, suggesting a much greater relaxation timescale when NO radiative cooling is not included in the calculation of thermospheric temperatures.

5.4 STUDY 3: OCTOBER 23-27th 2003

The ultimate test of the CMAT model's ability to correctly simulate the thermal response of the atmosphere to periods of high auroral activity involves modelling a real geomagnetic storm. During late October 2003, the Earth experienced a prolonged period of geomagnetic activity that included one of the largest geomagnetic storms ever measured. In this study, geomagnetic indices from the period 23rd October to 3rd November 2003 have been used to determine auroral energy inputs to the CMAT model. This period includes the 29th and 30th October, when the highest levels of activity were measured. A less intense, shorter lived geomagnetic storm occurred a few days prior to this on the 24th October, that was preceded by very low geomagnetic activity conditions, and followed by a reasonably uniform decrease in activity before the dramatic events of the 29th and 30th. A steady decrease in activity to moderate

conditions followed in the first few days of November. A CMAT simulation of this period will demonstrate the model's ability to respond to, and recover from a geomagnetic storm.

Temperature measurements made by the Fabry-Perot Interferometer at Kiruna in northern Sweden, have been used to validate the CMAT results. As a comparison, the CTIP GCM has been used to calculate temperatures over the same period. While the many differences between the two models make direct comparison difficult, it is nonetheless informative to see how the two models respond. Such an event is expected to result in the production of large quantities of NO and the inclusion of a self-consistent NO radiative cooling calculation in the CMAT model is expected to have a large impact on the simulated temperatures, over those calculated by CTIP.

5.4.1 SOLAR AND AURORAL ENERGY INPUTS

Using 3 hourly planetary K_p index values (taken from www.sec.noaa.gov) for the time period spanning 12:00 UT on 23rd October to 12:00 UT on 3rd November 2003, a time series of auroral power index (PI) has been derived. This power index is related to the planetary K_p by the statistical relationship of *Foster et al.* [1986], and determines the flux, energy and distribution of particle precipitation, along with patterns of high latitude convection electric fields used in the CMAT and CTIP models. K_p values measured over the 29th to 31st October exceeded those represented by the statistical patterns used in the models. For these levels of activity, precipitation and electric field patterns of the highest power index (10) have been used along with enhanced values of the total hemispheric precipitating power input. The values of hemispheric power used have been estimated based on the available data for lower levels of activity, as given by *Foster et al.* [1986]. In order to maintain model stability during prolonged periods of very high activity, the hemispheric power input has been capped at 150GW over the 29th and 30th October, and the CMAT model time step decreased from 60 to 10 seconds. CTIP has been run at its minimum time step of 30 seconds.

The solar flux input to the CMAT model has been varied each day, based on the F10.7 proxy for solar activity, also taken from www.sec.noaa.gov. In order to maintain

stability, the CTIP model has been run with constant solar inputs appropriate to an F10.7 of 150.

The time series of 3 hourly auroral power index and total hemispheric power inputs used in the CMAT and CTIP models are shown in Figure 5.10, along with F10.7 values used in the CMAT model.

5.4.2 NEUTRAL TEMPERATURE OVER KIRUNA

The Fabry-Perot Interferometer at the Kiruna Esrange Optical Platform System (KEOPS) in Sweden (67.8 N, 20.4 E) was used to measure the temperature of the neutral atmosphere throughout October and November 2003, when the Earth was subject to dramatic geomagnetic storms. By observing the red line aurora and airglow emission at 630nm wavelength, temperatures and winds at an altitude of around 240km can be derived. Neutral temperature measurements taken by the FPI at Kiruna between 12:00 UT 23rd October 3rd and 12:00 UT November 2003 are shown in Figure 5.10. Also shown are temperatures for approximately the same altitude, as calculated by the CMAT and CTIP models at 68° latitude, 20° longitude. Trends in measured temperatures clearly follow the changes in geomagnetic and solar activity, shown in the lower part of the figure. The match between variability in measured and calculated temperatures is good for both the CMAT and CTIP models, which predict rapid temperature changes in response to increases or decreases in geomagnetic activity. The most notable discrepancy occurring at times past 12:00 UT on 31st October, where a general downward trend in auroral energy inputs is accompanied by a continued increase in CTIP predicted temperatures. The timescale of recovery after the very high activity inputs of the 29th to 31st October is notably slower in the CTIP model than that predicted by the CMAT GCM and exhibited in the data.

Absolute magnitudes of observed and calculated temperatures are in good agreement over the majority of the study period. CMAT temperatures are consistently in better agreement with the data than those predicted by CTIP. This is in part due to the F10.7 values used in the CMAT model, which were changed each day in line with empirical values for those dates. To assess the impact of the F10.7 values used, a CMAT

simulation was performed covering six days from 23rd to 29th October using an F10.7 of 150, as was used in the CTIP runs. The CMAT temperatures were slightly lower than those shown in Figure 5.10 but still higher than those given by the CTIP model. The difference in temperatures calculated by the two models is therefore not a direct result of the level of solar activity specified. The different thermospheric heating schemes applied in each model will however have a large effect on simulated temperatures. CMAT uses high resolution solar flux data from the SOLAR2000 empirical model (*Tobiska et al.* [2000]), whereas CTIP uses a lower resolution spectrum based on that of *Hinteregger* [1970].

The very high temperatures measured by the FPI between 16:00 UT on 29th and 04:00 UT on 31st October are not attained by either model. This is the result of limitations in the statistical auroral energy inputs to the models, which do not account for the very high levels of geomagnetic activity encountered on those dates. Over periods of moderate to high geomagnetic activity, such as occurred between 18:00 UT on 25th and 06:00 on 29th October, the magnitude and variability of the CMAT predicted temperatures are in very good agreement with the FPI data.

The first large temperature increase measured by the FPI occurred on 24th October in response to a rapid increase in geomagnetic activity. On this date, the 3-hourly K_p index changed from 0^+ (00:00 UT to 03:00 UT) to 7^- (15:00 UT to 18:00 UT). Corresponding temperatures measured by the FPI are approximately 900K at 01:30 UT and 1200K at 16:30 UT. In the CMAT and CTIP models, this change in K_p is represented by a change in auroral power index from 1 to 10, and an increase in hemispheric power from 3GW to 135GW. Both the magnitude and increase in temperature predicted by the CMAT model compare well with the FPI data, increasing from around 825K at 01:30 UT to 1175K at 16:30 UT. CTIP temperatures change from approximately 870K to 1000K over the same period. After 20:00 UT, temperatures measured by the FPI continue to rise, reaching around 1350K at 04:40 UT on 25th October. The CTIP and CMAT temperatures however start to decrease at 22:00 UT and 00:00 UT respectively. This discrepancy between data and model arises from the statistical nature of the precipitation and electric field patterns in the models.

On 29th October, the three-hourly K_p index changed from 4°, for 03:00 UT to 06:00 UT, to 9⁻, for 06:00 UT to 09:00 UT. Geomagnetic activity remained at extremely high levels until 15:00 UT on 31st October, after which a reasonably steady decrease to moderate activity occurred. Temperatures measured by the FPI increased dramatically from around 1100K at 04:00 UT to over 2000K at 20:00 UT on 29th October. At 16:00 UT on 31st October, the observed temperature had relaxed to around 1200K. As activity levels continued to decrease, so too did the observed temperature, which reached around 950K at 00:00 on 3rd November. As previously mentioned, the statistical patterns of particle precipitation and convection fields used in the CMAT and CTIP models are not able to accurately represent the extremely high levels of auroral activity observed from 29th to 31st October 2003. The absolute magnitudes of temperatures predicted by the models for this period are therefore somewhat lower than those seen in the data. Maximum temperatures predicted by the CMAT and CTIP models for the 29th October were 1500K and 1260K respectively. The models are able to recreate the rapid increases in temperature associated with the high auroral forcing, although the CTIP model does lag behind CMAT, in both the maximum response and recovery timescales. The realistic post storm thermal recovery predicted by the CMAT model is thought to be a direct result of NO cooling, which is calculated in the CMAT model but absent from the CTIP calculations.

5.4.3 MERIDIONAL VARIATION IN THERMAL RESPONSE

The neutral temperature at approximately 250km calculated by the CMAT model is shown as a function of latitude and time in Figure 5.11, for the period 12:00 UT on 23rd October to 12:00 UT on 3rd November 2003. Plots are shown for 0° and 180° longitude. The distribution of energy deposition will be different in each longitude sector, depending on the convection and particle precipitation patterns used. A detailed comparison of temperatures in the two plots would therefore be misleading. A qualitative view of the meridional extent of temperature enhancements in the two sectors can however be gained from each plot.

As might be expected, temperatures are generally greater in the southern hemisphere which is exposed to greater levels of solar radiation than the northern hemisphere at

this time of year. Joule heating leads to high latitude temperature enhancements in both hemispheres. The level of Joule heating is directly affected by changes in geomagnetic activity and so therefore, is the temperature of the high latitude thermosphere. Increases of up to 450K and 500K occur in the high latitude northern and southern hemispheres respectively in response to the dramatic changes in geomagnetic activity on 24th October. In the northern hemisphere, the aurorally induced temperature changes are fairly localised and do not extend equatorward of about 60° latitude. In the southern hemisphere, the offset between the geographic and geomagnetic poles is greater, meaning aurorally induced temperature enhancements extend further towards the equator, to around -40° latitude. Where increased Joule heating occurs in the day-lit sector of the southern hemisphere, temperature enhancements are seen to extend all the way from the pole to the equator. This occurs as low and mid latitude locations undergo heating though the absorption of solar radiation.

From 29th to 31st October geomagnetic activity reached extremely high levels. CMAT temperatures in the southern hemisphere high latitude region reach up to 1800K. In the northern hemisphere, temperatures of just under 1600K are predicted. Very high levels of solar heating result from the large F10.7 values used to determine solar flux inputs to the model. As a result, equatorial temperatures reach over 1600K. At 180° longitude in the southern hemisphere, aurorally induced temperature enhancements reach within 10° latitude of the equator in the early hours of 1st November. While expansion of the auroral oval will contribute to the equatorward spread of Joule heating, the large temperatures predicted at these low latitudes are almost certainly the result of transport. Strong meridional winds are induced by pressure gradients that result from Joule heating. These winds act to transport energy and chemical constituents to lower latitudes.

5.4.4 EVOLUTION OF NITRIC OXIDE DISTRIBUTION

Figure 5.12 shows the CMAT calculated Nitric oxide number densities at approximately 110km altitude as a function of latitude and time, for 0° and 180° longitude. The time series covers the period from 12:00 UT on 23rd October to 12:00 UT on 3rd November 2003. Note the difference in scale between the plot for 0° and

180° longitude. The concentrations of NO at high latitudes are greater than those at low latitudes. Peak densities in the northern hemisphere occur closer to the pole than those on the south. This is a result of the increased offset between geographic and geomagnetic poles in the southern hemisphere. The NO abundance clearly responds to changes in geomagnetic activity. At 0° longitude, maximum concentrations of just under 9.0 x $10^{14}m^{-3}$ occur in the southern hemisphere between –30° and –40° latitude, around midday on the 30th and 31st October. Maxima of nearly 14 x $10^{14}m^{-3}$ are predicted for the southern hemisphere at 180° longitude, between –10° and –30° latitude. The maxima occur between 12:00 UT and 22:00 UT. While the greatest increases in temperature occur very rapidly after changes in geomagnetic activity, the maximum NO concentrations occur one to two days after the onset of a period of high activity.

For the time period modelled here, the high latitude regions of the northern hemisphere are less affected by solar radiation than equivalent latitudes in the southern hemisphere. This means the effects of aurorally produced NO can be more easily recognised than in the southern hemisphere where the sources of NO production are indistinguishable. The latitudes at which NO maxima occur in the northern hemisphere range between 80° on the 28th October, to 50° on the 31st October. The peak clearly moves towards the equator as geomagnetic activity is increased. Expansion of the auroral oval will contribute to this equatorward progression, as will transport of NO by strong meridional winds.

At mid and low latitudes, the calculated concentrations of NO increase dramatically over the simulation period. Figure 5.13 shows the equatorial NO number density at approximately 110km altitude for 0° and 180° longitude. The diurnal variation in NO density results from the changing flux of solar soft X-rays over the day. At 0° longitude this variation dominates from 12:00 UT on 23rd to the early hours of 1st November, after which a significant contribution from aurorally produced NO acts to alter the diurnal structure. The gradual increase in daily NO concentrations seen between the 23rd and 30th October results from increases in the solar flux input to the model, as determined by the F10.7. At around 22:00 UT on 23rd and 29th October,

peaks in the NO density of 1.5 and 2.9 $10^{14}m^{-3}$ are seen. The F10.7 values used on these dates are 181 and 275 respectively. On 1st November, a peak NO number density of just under 3.3 $10^{14}m^{-3}$ is predicted to occur at around 08:00 UT. The F10.7 used for this day was 245. The fact that this large NO peak is created at 08:00 UT on a day when the trend in F10.7 is downward indicates that there is a significant contribution from aurorally produced NO.

At 180° longitude, the equatorial NO density also displays a strong diurnal variability from 23rd to 29th October, with peak NO densities of 1.5 and 2.6 $x10^{14}m^{-3}$ occurring close to 10:00 UT on the 24th and 29th October respectively. In the early hours of 30th October, a large increase in NO density is predicted, approximately 20 hours after the start of the very intense storm period. By 12:00 UT on 1st November, the NO density has reached a peak of 11.6 $x10^{14}m^{-3}$. This very high number density arises as aurorally produced NO is transported from high latitudes. While the equatorial NO density does start to drop after this peak, large concentrations are still present at the end of the simulation period.

5.5 CONCLUSIONS

Three studies have been performed using the CMAT GCM. The first explores the latitudinal extent and transport of aurorally produced NO under stable conditions of moderate geomagnetic activity. The second investigates the response of modelled temperature and NO densities to a period of high geomagnetic activity. In the third, the CMAT model is used to simulate an 11 day period from 23rd October to 3rd November 2003 during which extremely high levels of geomagnetic activity were recorded. The conclusions from each study are given below.

5.5.1 STUDY 1

In the first study, three CMAT simulations were performed, one with no auroral energy inputs, one with inputs appropriate to moderate geomagnetic activity, and one with moderate activity auroral energy inputs but no horizontal or vertical advection. Comparison of Nitric oxide number density profiles resulting from these simulations shows that during conditions of moderate geomagnetic activity, high latitude nitric

oxide densities are enhanced over those at low latitudes. At 0° longitude, maximum NO enhancements of up to 2.5 x 10^{14} m^{-3} occurred at approximately ±80° latitude, 110km altitude, depending on time of day. Over an entire day, at 110km the most equatorward latitudes at which aurorally produced NO occurred were -30° in the southern hemisphere and 45° in the north.

While variations in atmospheric dynamics will strongly affect the nature and extent of constituent transport, for the conditions simulated here, horizontal advection acts to transport NO from the high to low latitude regions. When advective terms are included in the minor constituent continuity equation, NO densities at mid-latitudes are enhanced by varying amounts, depending on location and time of day. In the southern hemisphere where the offset between the geomagnetic and geographic pole is greatest, NO can be transported to within 30° of the equator. At low latitudes, NO densities are strongly influenced by the vertical diurnal tide.

5.5.2 STUDY 2

The second study involved two CMAT simulations, both for a period of 6 model days. The first simulation included a 14-hour period of enhanced geomagnetic activity, while the second used constant moderate activity inputs throughout the entire 6 days. Comparison of the NO densities at 110km calculated in each model run suggests that the increases in NO density resulting from a period of enhanced auroral activity are highly variable in spatial and temporal distribution. The magnitude of NO enhancements at a particular geographic point are dependent on location with respect to the auroral oval. Satellite measurements of NO density have revealed a 1-day lag in the response of NO to high latitude forcing (e.g. *Marsh et al.* [2004], *Barth and Bailey* [2004]). The results of these simulations confirm that there is a lag in the maximum response of NO density, but the size of the lag varies with geographic location in relation to the auroral oval. Mid latitude NO maxima can occur between 14 and 48 hours after the end of the storm forcing period. The local time at which auroral forcing occurs also affects the size of NO enhancements and their transport to lower latitudes.

In the simulations performed here, mid and high latitude NO enhancements remain significant for several days after relaxation of a storm. The lifetime of NO enhancements produced by auroral energy input is dependent on the number of hours for which the enhanced region is exposed to sunlight. Season and latitude are therefore key factors in determining the longevity of aurorally produced NO.

Investigation of the thermal response to storm forcing again reveals a complex pattern of temperature enhancements, dependent on latitude, longitude and time of day. High latitude temperature increases occur rapidly after storm onset then saturate around 8 to 12 hours into the high activity period. At 130km altitude maximum temperatures of 800K are predicted for high latitude locations, while low latitude temperatures reach 580K. At 300km maximum temperatures of 1300K and 1075K are simulated for high and low latitude locations respectively. Temperature maxima occur between 4 and 16 hours after storm onset depending on geographic location.

A comparison of the control and storm simulations reveals high latitude temperature enhancements of up to 250K at 130km altitude when the period of high activity is included. At 300km, high latitude temperature enhancements reach a maximum of 370K. Low latitudes undergo maximum temperatures increases of 50K and 80K at 130km and 300km respectively. The spatial and temporal distribution of temperature enhancements is again highly variable with geographic location. High latitude temperature enhancements are greater in magnitude and longer lived than at low latitudes. At both 130 and 300km altitude, sharp temporal gradients in temperature result from rapid changes in auroral activity.

Comparison of zonally averaged temperatures in the control and storm simulations reveals a relatively smooth relaxation to pre-storm temperatures at 120km, over a timescale of approximately 60 hours after the end of the storm. At higher altitudes, temperatures reach their pre-storm values just 36 hours after the end of the high activity period. A small oscillation about the pre-storm value is seen. An identical simulation by the CTIP model shows considerably greater relaxation timescales that extend beyond the end of the simulation period at all altitudes tested (120 to 300km).

The difference between the models is attributed to NO cooling which is not accounted for in the CTIP model, but is self consistently calculated in CMAT.

5.5.3 STUDY 3

In the final modelling study, the CMAT and CTIP models have been used to simulate the period from 12:00 UT on 23rd October to 12:00 UT on 3rd November 2003, during which the Earth experienced some of the largest geomagnetic storms ever measured. Neutral temperatures calculated by the CMAT and CTIP models were compared with FPI measurements made in Kiruna in northern Sweden. The magnitude and variability of the CMAT temperatures matched the FPI data well, although the very high temperatures recorded at the peak of the largest storm were underestimated by the model. CTIP predicted temperatures that were slightly lower than those of CMAT and the FPI data, and the timescales of post-storm recovery were overestimated. The improved temperature variability modelled by CMAT is attributed to the inclusion of self-consistent nitric oxide cooling. Differences in the magnitude of simulated temperatures are thought to be due primarily to differences in the thermospheric heating routines used in each model, associated with the solar flux inputs used.

The meridional distribution of temperature enhancements resulting from the geomagnetic storm at approximately 250km altitude has been assessed. For the dates in question, high latitudes in the northern hemisphere are less influenced by solar radiation than the equivalent latitudes in the south. Temperatures are therefore generally greater in the south than in the north. High latitude temperatures respond to changes in geomagnetic activity as the associated level of Joule heating changes. In the northern hemisphere, temperature changes associated with the rapid increase in geomagnetic activity on 24th October are restricted to latitudes poleward of about 60° latitude. In the southern hemisphere, aurorally induced temperature enhancements remain at latitudes poleward of about -40° latitude. During the period of maximum activity from 29th to 31st October, CMAT temperatures reached up to 1800K and 1600K in the southern and northern hemispheres respectively. Aurorally induced temperature enhancements reached 10° south of the equator in the early hours of November 1st. This is the result of strong meridional winds associated with high levels

of Joule heating. Increases in the F10.7 solar activity proxy throughout the first 7 days of the simulation lead to enhanced equatorial temperatures of up to 1600K.

Nitric oxide number densities at the peak altitude of 110km respond to changes in the solar and geomagnetic energy inputs. The magnitude and location of NO density peaks are highly variable, depending on location with respect to the auroral oval, transport of NO by meridional winds, and changes in solar flux inputs. The latitude at which maxima occur moves towards the equator as geomagnetic activity increases and the auroral oval expands. Maximum concentrations modelled by CMAT reach just less than 14 $x10^{14}m^{-3}$ and occur between –10° and –30° latitude, 180° longitude in the southern hemisphere. The peaks in NO concentrations occur one to two days after the onset of a period of high activity.

Prior to the large increases in geomagnetic activity seen on the 29th October, equatorial NO densities are largely controlled by diurnal variations in the solar energy input. Increases in the equatorial NO budget start to occur 1 day after the onset of the storm, depending on location. These enhancements arise as NO produced at high latitudes is transported equatorward by strong meridional winds. Maximum equatorial NO densities of 3.3 and 11.6 $x10^{14}m^{-3}$ are predicted to occur at 08:00 UT and 12:00 UT on 1st November, at longitudes of 0° and 180° respectively.

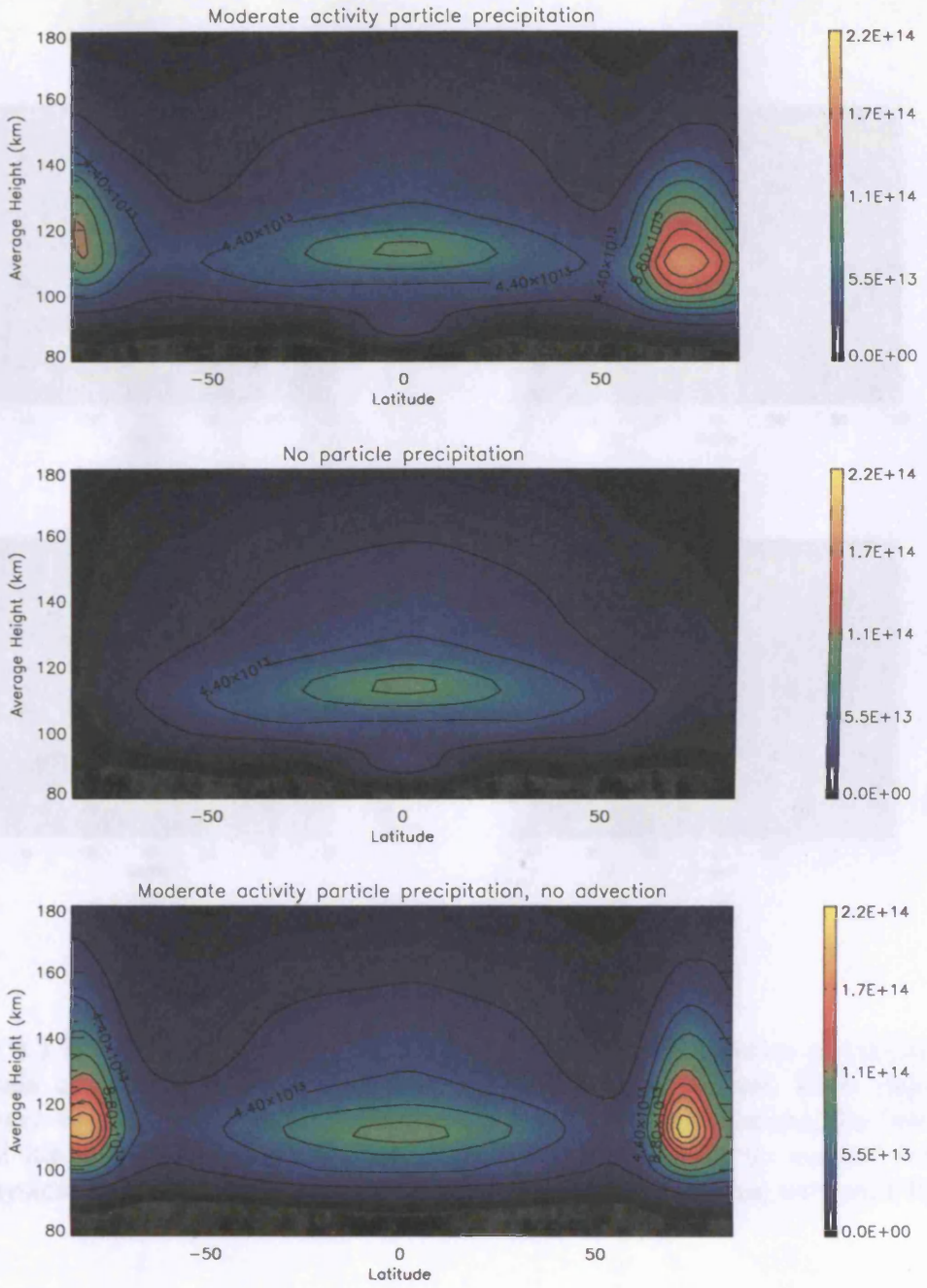


Figure 5.1 Latitude vs. height plots of Nitric Oxide number density (m^{-3}) at a longitude of 0°, at 12:00 UT. Profiles are as calculated by CMAT with auroral energy inputs appropriate to a K_p of 2^+ (*top*), with no auroral inputs (*middle*), and with auroral energy inputs but no advection (*bottom*). Conditions are appropriate to northern spring equinox at an F10.7 of 105.

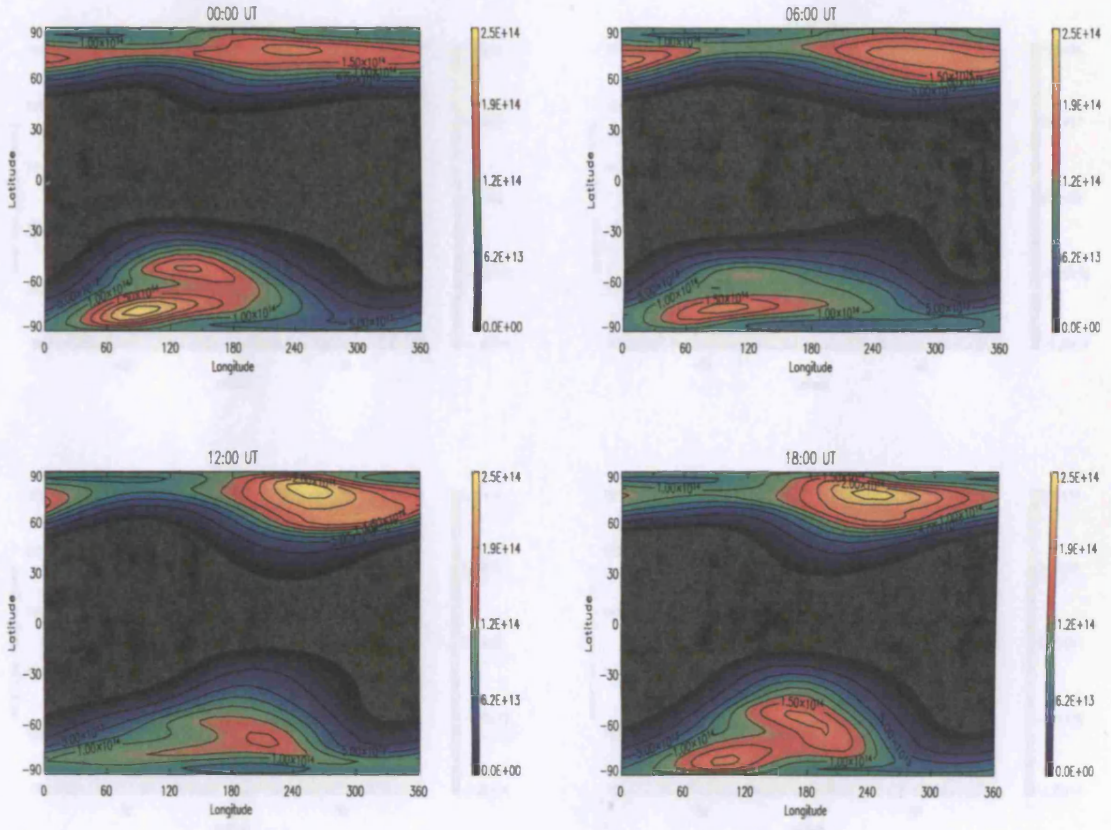


Figure 5.2 Difference plots of NO number density (m^{-3}) as a function of latitude and longitude at an altitude of 110km at 4 different universal times. Plots show the difference between NO densities calculated in two CMAT simulations, the first with auroral inputs appropriate to a K_p of 2^+ and the second with no auroral forcing. Geophysical conditions are appropriate to northern spring equinox with an F10.7 of 105.

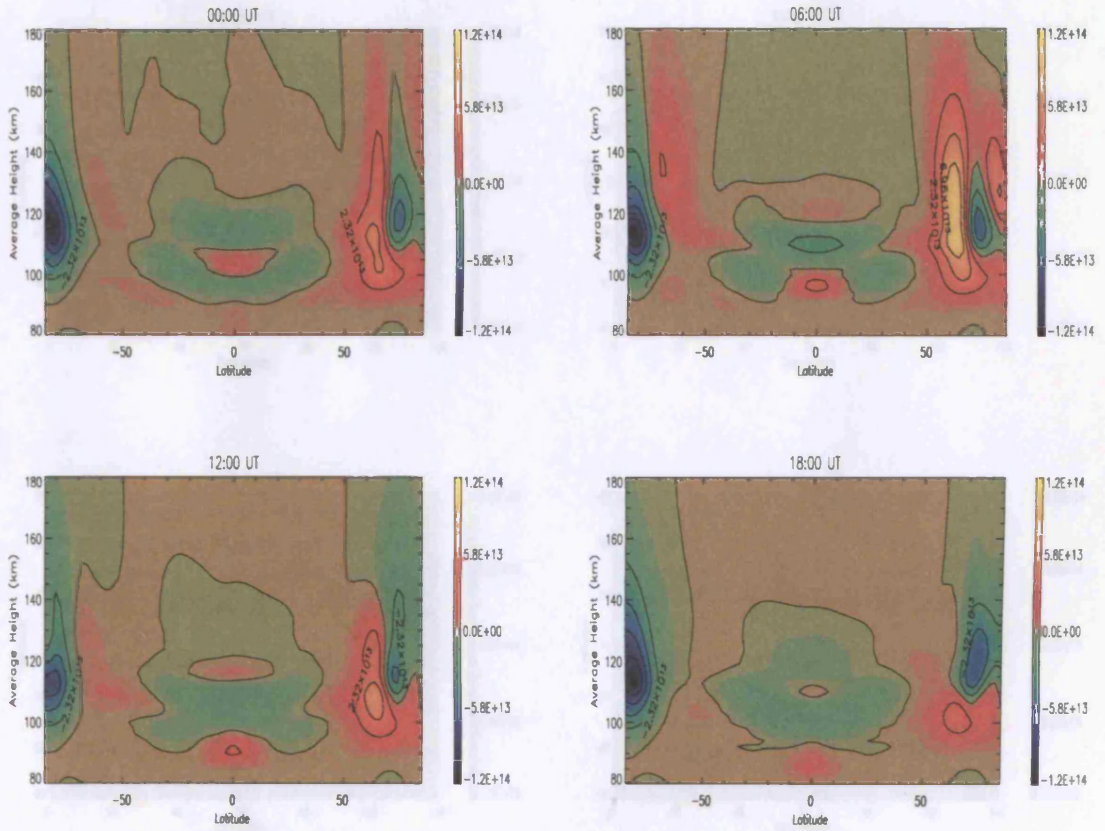


Figure 5.3 Difference plots of NO number density (m^{-3}) at 0° longitude as a function of latitude and altitude at 4 different universal times. Plots show the difference between NO densities calculated in two CMAT simulations, the first with minor species advection included in the NO density calculation, the second with no transport terms. Geophysical conditions are appropriate to northern spring equinox with an F10.7 of 105.

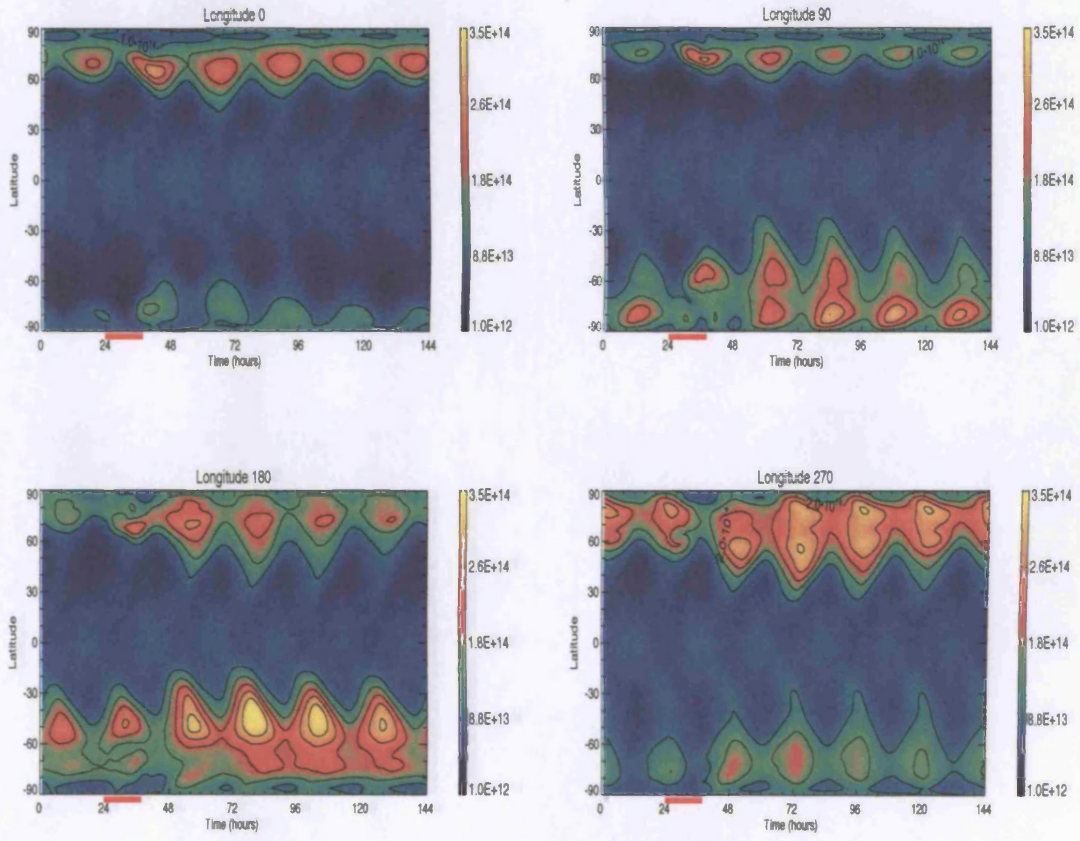


Figure 5.4 Nitric oxide number density (m^{-3}) at approximately 110km altitude as a function of time and latitude at 0°, 90°, 180° and 270° geographic longitude. The time axis refers to hours from the start of a CMAT run that includes high activity auroral forcing between 24:00 and 38:00 hours, as indicated by the red line. Auroral forcing before and after this period is characteristic of moderate geomagnetic activity. Conditions are appropriate to northern spring equinox with an F10.7 of 105.

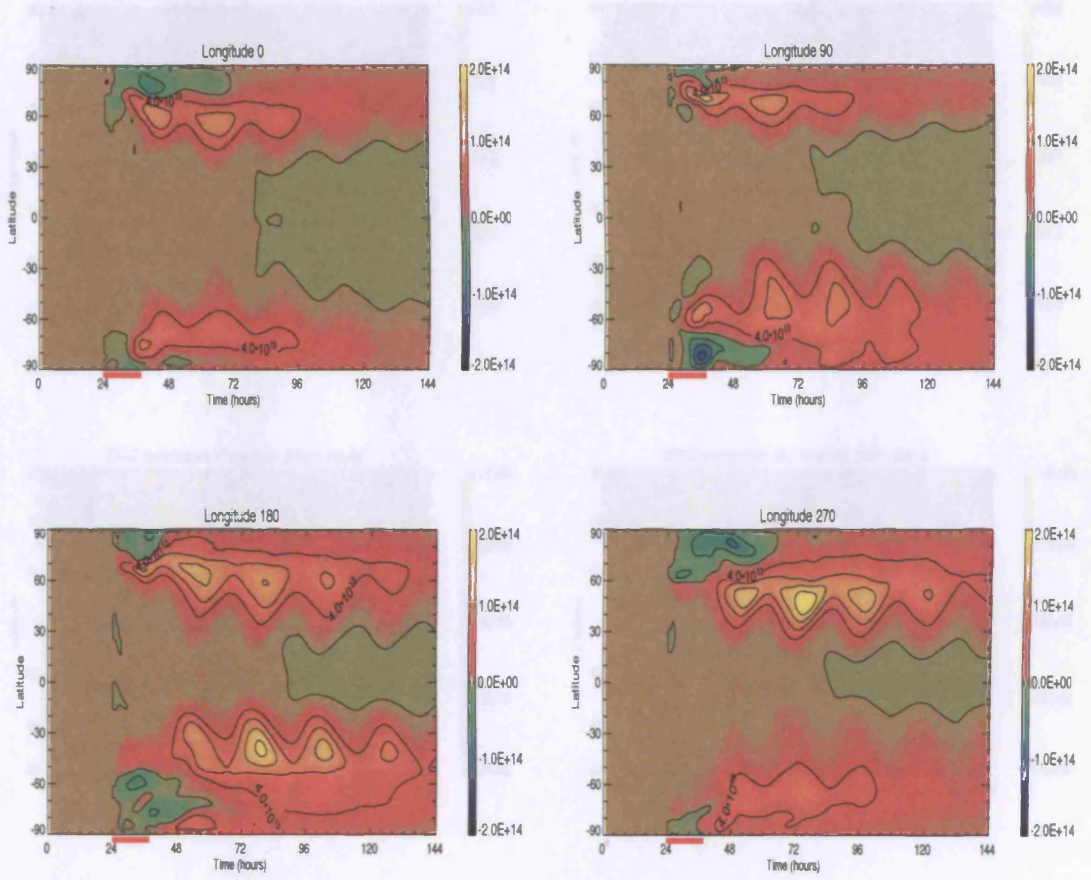


Figure 5.5 Difference plots of Nitric Oxide number density (m^{-3}) at approximately 110km altitude, at 0°, 90°, 180° and 270° geographic longitude. NO densities were calculated in two CMAT runs, the first with high activity auroral forcing between 24:00 and 38:00 hours (as indicated by the red line) and moderate forcing for the remaining hour; the second with constant moderate activity forcing throughout the full 5 days of the simulation. Conditions are appropriate to northern spring equinox with an F10.7 of 105.

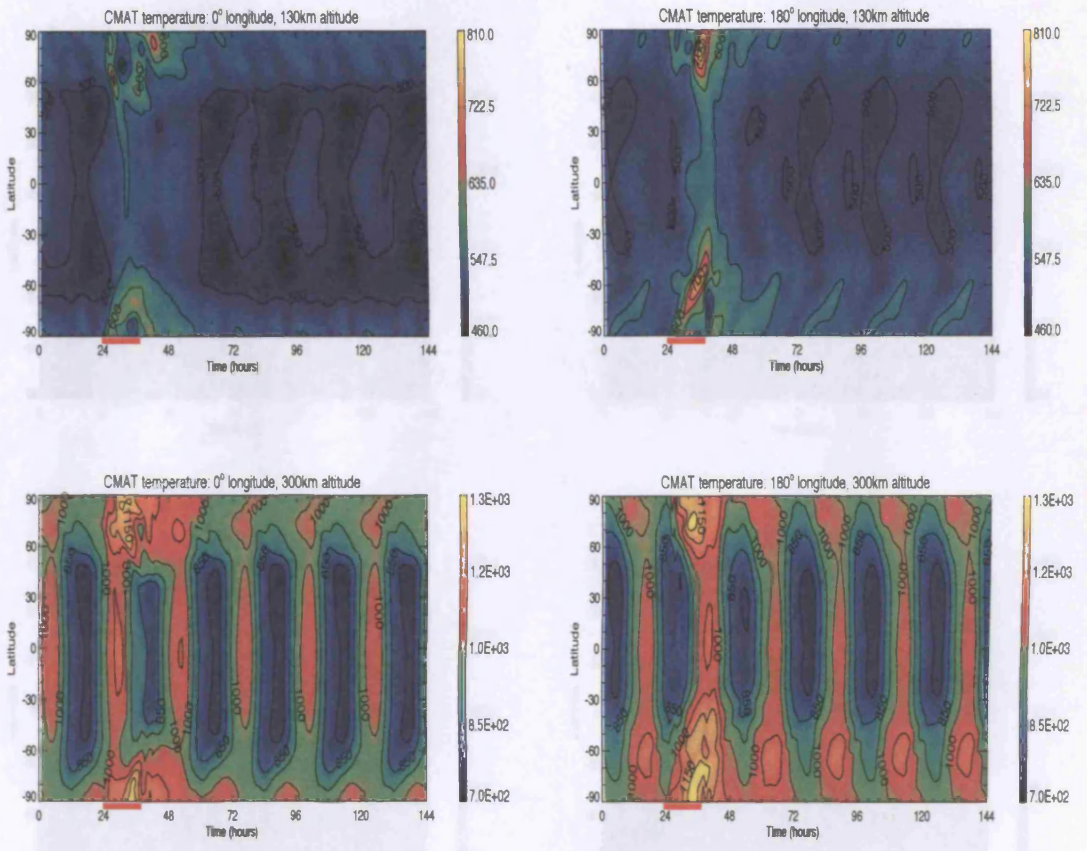


Figure 5.6 Neutral temperature (K) calculated by the CMAT model as a function of geographic latitude and time at approximately 130 (*top*) and 300km (*bottom*) altitude, at 0° (*left*) and 180° (*right*) geographic longitude. The time axis refers to hours from the start of the 6 day simulation where high energy auroral forcing has been applied between hours 24:00 and 38:00, as indicated by the red line.

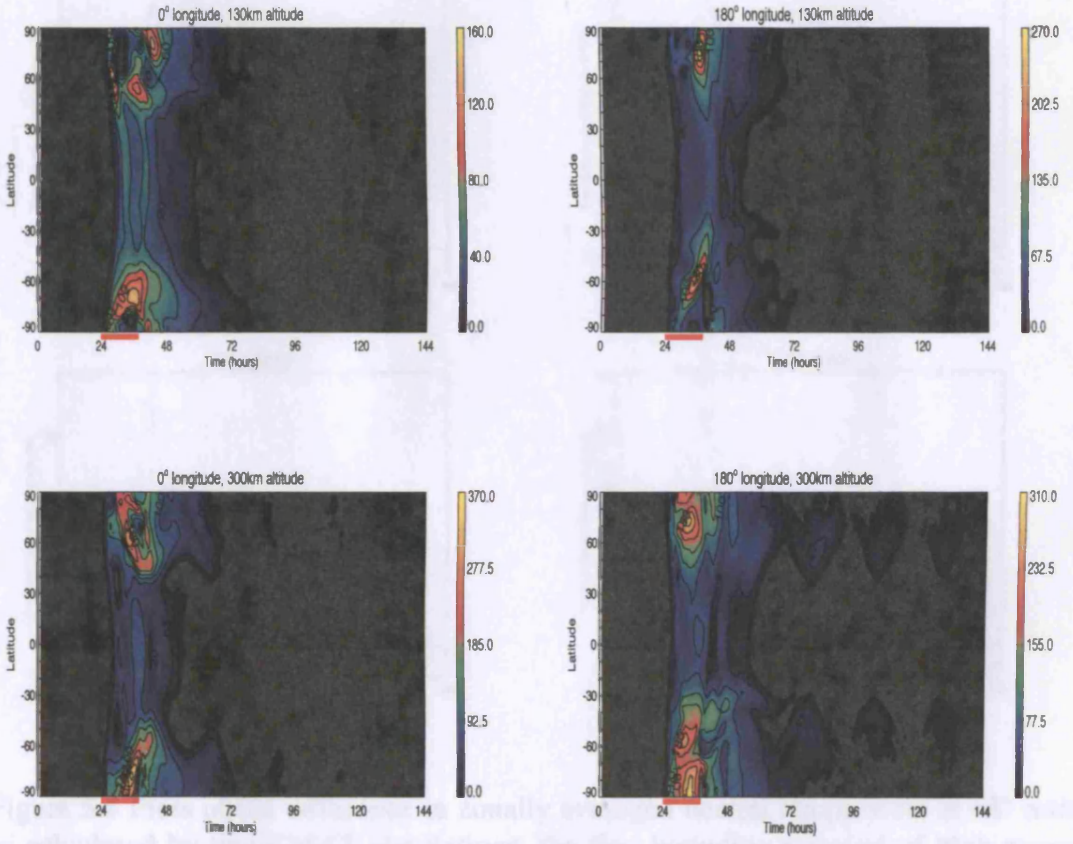


Figure 5.7 Calculated neutral temperature difference between two CMAT simulations, the first including a period of high energy auroral forcing between hours 24:00 and 38:00 of the 6 day run (as indicated by the red line), the second using auroral forcing appropriate to moderate geomagnetic activity throughout the full 6 days. Data is shown for altitudes of approximately 130 (*top*) and 300km (*bottom*) at 0° (*right*) and 180° (*left*) geographic longitude. Note the change in scale between the plots at 130km altitude.

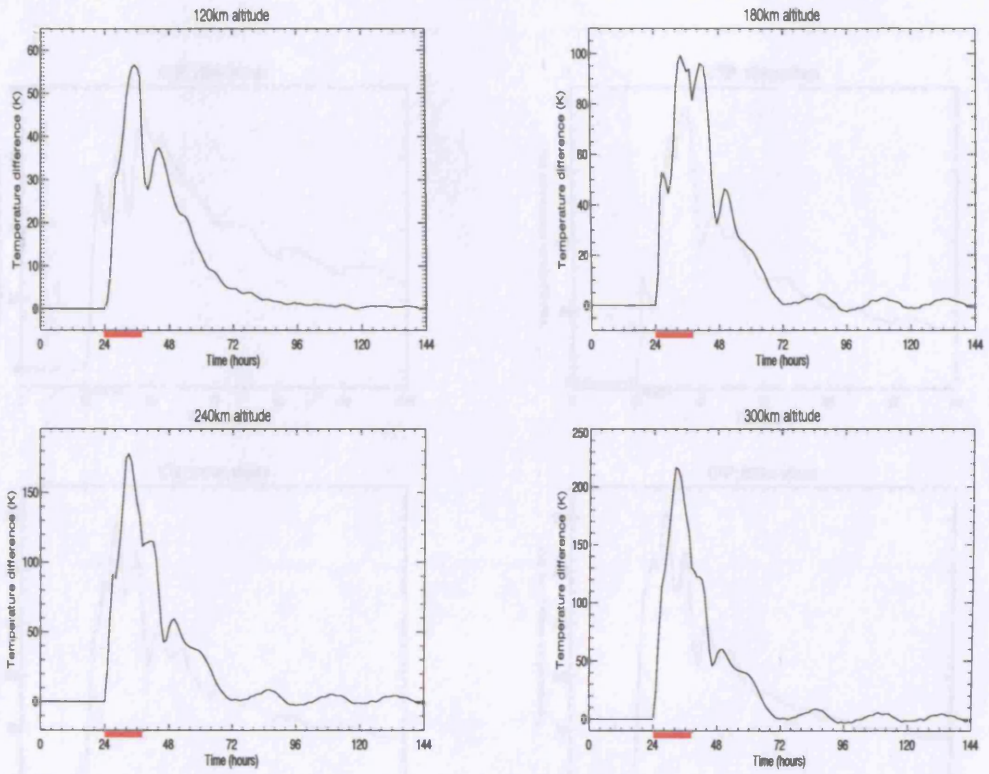


Figure 5.8 Plots of the difference in zonally averaged neutral temperature at 65° north as calculated by two CMAT simulations, the first including a period of high energy auroral forcing between hours 24:00 and 38:00 of the 6 day run (as indicated by the red line), the second using auroral forcing appropriate to moderate geomagnetic activity throughout the full 6 days. Data is from altitudes of approximately 120, 180, 240 and 300km. The time axis refers to hours from the start of the model run.

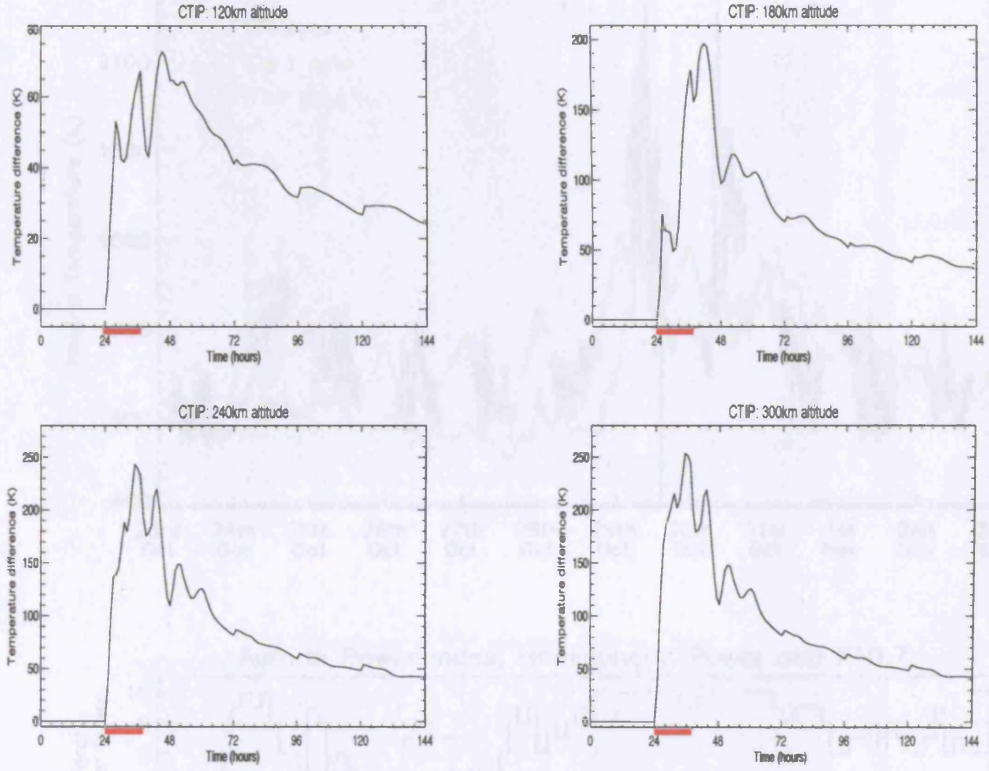


Figure 5.9 As for Figure 5.8 but using zonally averaged temperature data from the CTIP model.

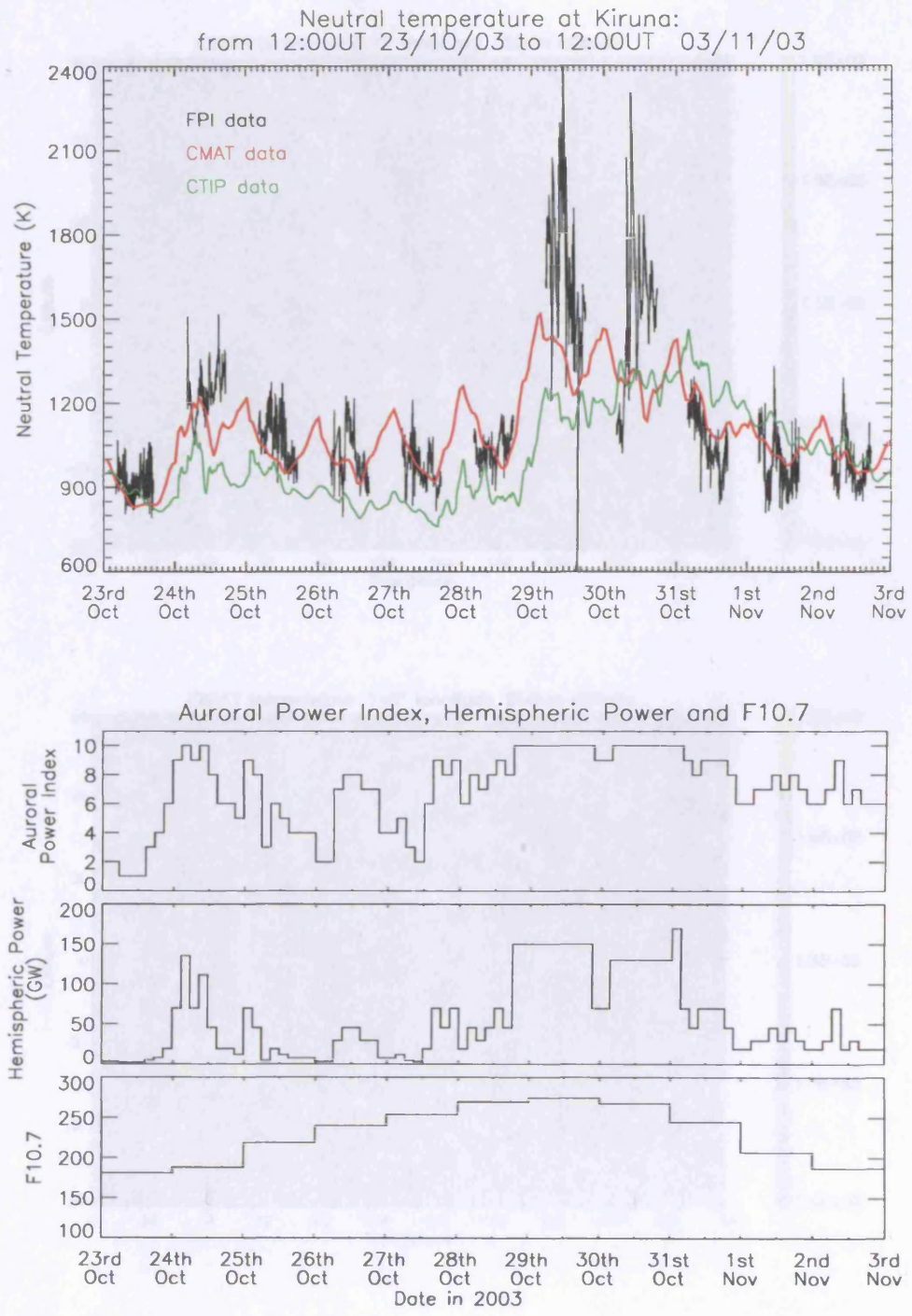


Figure 5.10 Neutral temperature at 250km altitude above Kiruna, northern Sweden, from 12:00UT on 23rd October to 12:00UT on 3rd November 2003, as measured by FPI and as calculated by the CMAT and CTIP models (*top*). Auroral power index and hemispheric power inputs used in the models for the same time period, along with the associated F10.7 inputs to the CMAT model (*bottom*).

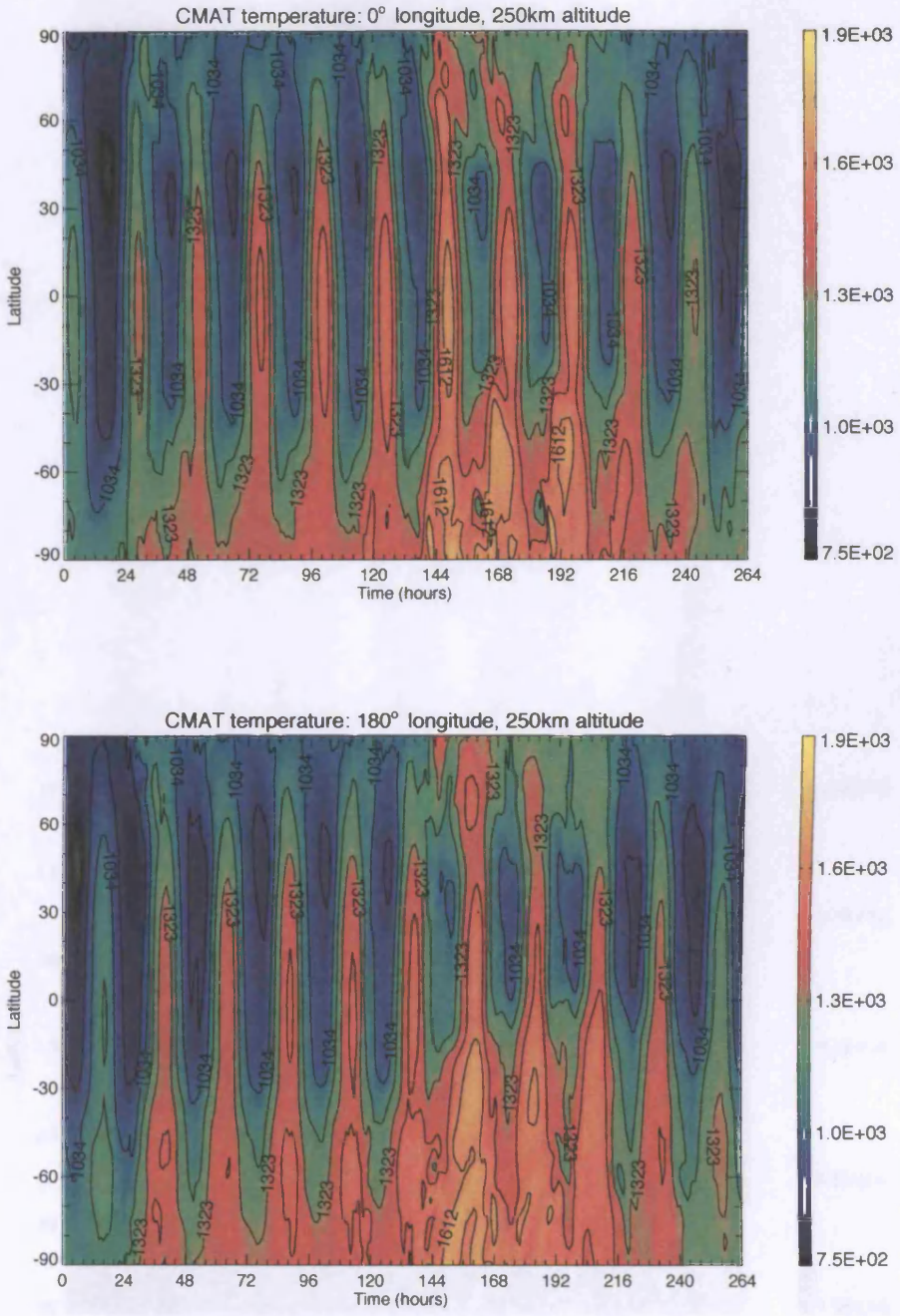


Figure 5.11 CMAT calculated neutral temperature at 250km altitude as a function of latitude and time at 0° (*top*) and 180° (*bottom*) longitude. The time axis covers the period from 12:00UT 23rd October to 12:00UT 3rd November 2003.

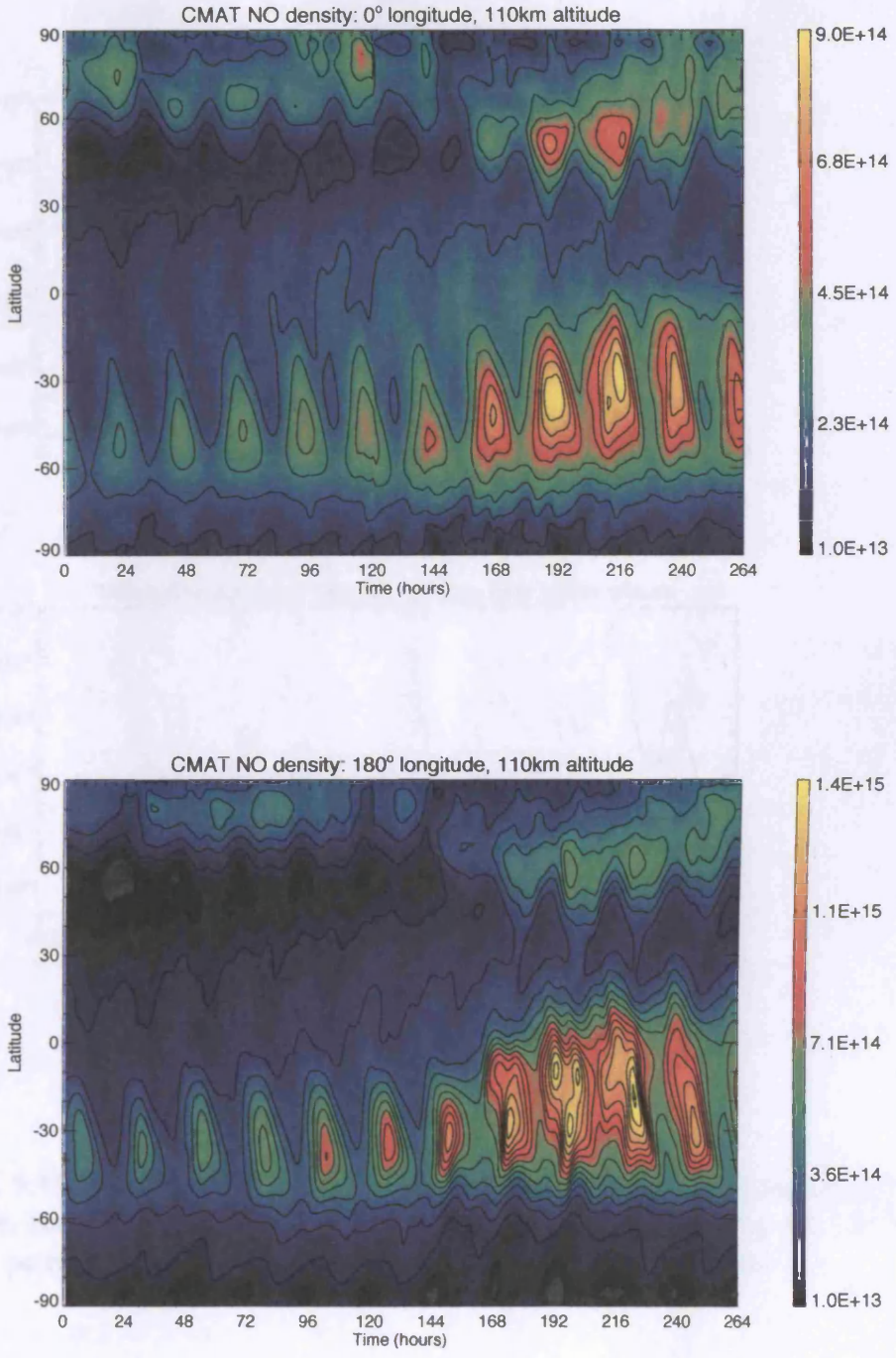


Figure 5.12 CMAT calculated nitric oxide number density (m^{-3}) at approximately 110km altitude as a function of latitude and time, at 0° (*top*) and 180° (*bottom*) longitude. The time axis covers the period from 12:00UT 23rd October to 12:00UT 3rd November 2003. Contours are every $1x10^{14}$.

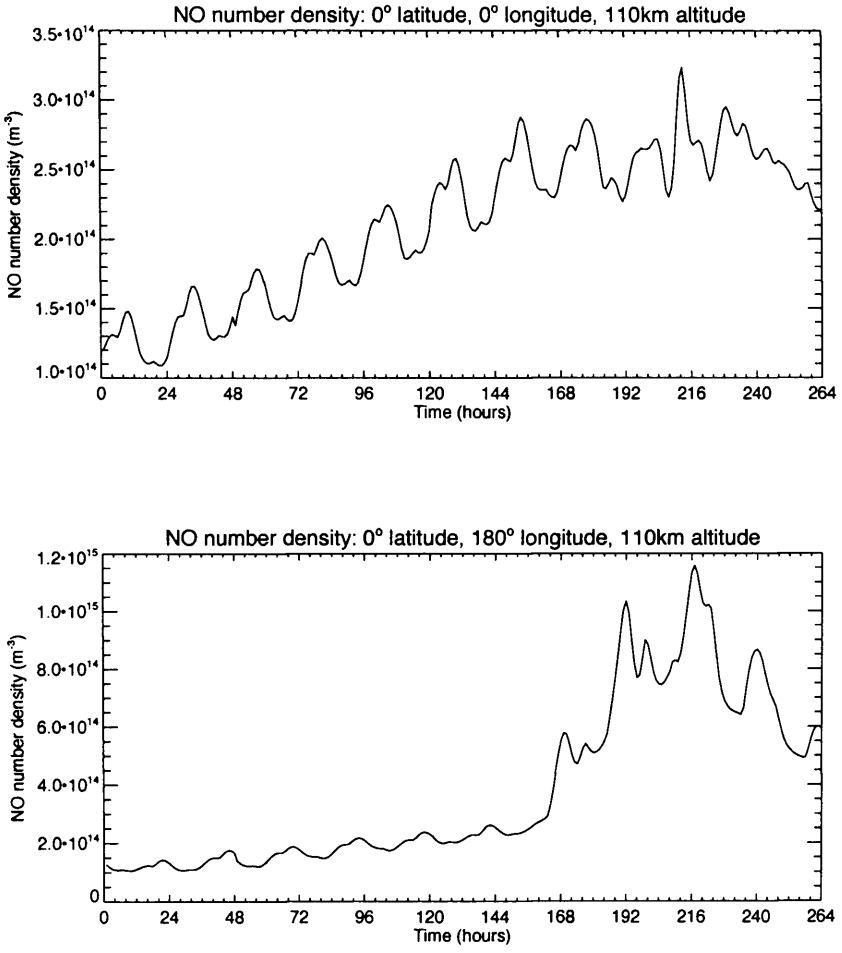


Figure 5.13 Equatorial nitric oxide number density (m^{-3}) at approximately 110km altitude, for 0° (*top*) and 180° longitude (*bottom*), as calculated by the CMAT model for the period 12:00UT 23rd October to 12:00UT 3rd November 2003.

CHAPTER VI. THE EFFECT OF GRAVITY WAVE DRAG PARAMETERISATIONS ON THE DIURNAL TIDE IN CMAT. IMPLICATIONS FOR THE O(^{1}S) GREEN LINE VOLUME EMISSION RATES IN THE MLT REGION.

6.1 INTRODUCTION

Non-linear coupling between gravity waves and tidally induced temperature and wind perturbations can lead to complex interactions between energetics and dynamics which can in turn, have a significant impact on atmospheric composition. *Fritts and Vincent* [1987] were amongst the first to suggest that observed tidal variability could be attributed to gravity wave-tidal interactions when they identified distinct diurnal variation in MF radar observations of gravity wave momentum flux. They also suggested that momentum flux associated with gravity wave breaking could act to advance the tidal phase and decrease the amplitude. Several modelling studies aimed at exploring interactions between gravity waves and the diurnal tide have also implied that gravity wave breaking acts to damp the diurnal tide (*Forbes et al.* [1991], *Miyahara and Forbes* [1991], *Meyer* [1999a]). In recent years it has however become clear that the representation of gravity wave breaking in atmospheric models has a large impact on the amplitude and phase of the simulated diurnal tide. This means the implementation of different gravity wave schemes in numerical models can result in significantly different tidal winds.

McLandress [1998] compared changes in the diurnal tide resulting from the implementation of three commonly used gravity wave parameterisations in a zonal mean circulation model. Two of the parameterisations resulted in a decrease in tidal amplitude, namely that of *Fritts and Lu* [1993], and a scheme based on *Lindzen* [1981]. Conversely the *Hines* [1997a,b] Doppler spread parameterisation caused an increase in the amplitude. A similar result was reported by *Mayr et al.* [2001] who noted an increase in the diurnal tidal amplitude when using the *Hines* scheme in a 3D numerical model. *Akmaev* [2001] argued that the level of tidal damping observed in the results of *Miyahara and Forbes* [1991] was overestimated due to incorrect

implementation of the *Lindzen* [1981] scheme used. All the schemes tested by *McLandress* [1998] acted to advance the phase of the tide such that when gravity wave drag was included, the peak amplitude at a given height occurred at an earlier time of day. While it is accepted that an increase in turbulent eddy mixing due to gravity wave breaking results in a dampening of the tide, the role of gravity wave momentum deposition is still uncertain (*Meyer* [1999a], *Akmaev* [2001]).

Results from the Canadian Middle Atmosphere Model (CMAM) (*Beagley et al.* [1997]) using the gravity wave parameterisations of *Hines* [1997a,b] and of *Medvedev and Klaassen* [2000] have been compared with observations from MF radars at different latitudes (*Manson et al.* [2002]). Both parameterisations acted to modify the amplitude and phase of the tides by varying amounts, depending on season, latitude and tidal mode. Tidal phases modelled using the *Medvedev and Klaassen* scheme were generally in better agreement with low latitude MF radar observations than those modelled using the *Hines* parameterisation, as were the mean zonal wind fields. One should note however that many of the results presented were either at an altitude where tidal amplitudes are relatively small (78km), or close to the upper boundary of the model[ID280]. The authors suggested that the effect of gravity wave parameterisations on tides could be better assessed at a height of 90km. What is clear is that the choice of gravity wave parameterisation in modelling studies can have a significant influence on the results. Further observational and modelling work is therefore needed to better appreciate the effects of parameterisations in models and their ability to represent complex gravity wave-tidal interactions in the atmosphere.

Observed O(^{1}S) 557.7nm green line volume emission rates in the MLT region have been shown to be strongly influenced by the migrating diurnal tide (*McLandress* [1994], *Shepherd et al.* [1995], *Yudin* [1998], *Roble and Shepherd* [1997]). Modelling studies of this airglow can be used to interpret the dynamical influence of the tide on emission rates. Valuable information about the distribution of associated species can also be gained. The modelling studies of *Roble and Shepherd* [1997], *Yee et al.* [1997] and *Harris et al.* [2002], all used global GCMs to study the local time variation of the low latitude O(^{1}S) emission rate. Strong lower boundary tidal forcing was required in

order to reasonably reproduce the diurnal variability in airglow that has been observed by the Wind Imaging Interferometer (WINDII) instrument on the Upper Atmosphere Research Satellite (UARS) (e.g. *Shepherd et al.* [1995]). Increased lower boundary forcing also improved agreement between meridional winds calculated by the CMAT model and those observed by the High Resolution Doppler Imager (HRDI) on board the UARS satellite. *Harris et al.* [2002] suggested that the increased tidal forcing was probably required as tides (taken from the Global Scale Wave Model (*Hagan et al.* [1995])) applied at the lower boundary of the model, were being damped by the hybrid *Lindzen-Matsuno* gravity wave scheme used (*Meyer* [1999a,b]). The studies by *Roble and Shepherd* [1997] and *Yee et al.* [1997] used the gravity wave parameterisation of *Fritts and Lu* [1993] in the National Center for Atmospheric Research Thermosphere-Ionosphere-Mesosphere-Electrodynamics General Circulation model (NCAR TIME-GCM, *Roble and Ridley* [1994]). This scheme has also been shown to damp the tide (*McLandress* [1998]). It therefore seems likely that the choice of gravity wave parameterisation in the CMAT and TIME GCMs has an effect on the modelled meridional winds, tidal amplitudes, atomic oxygen concentration and associated green line volume emission rate.

Both *Harris et al.* [2002] and *Roble and Shepherd* [1997] speculated that an overestimation of eddy diffusion in the GCMs may result in an increase in O recombination. This was proposed as the reason for the factor of 2 deficit in modelled green line emission rates in comparison to those measured by WINDII. An overestimation of eddy diffusion in the models would also contribute to a decrease in tidal amplitude (*Meyer* [1999a]).

Zhang et al. [2001] found they were able to gain good agreement between the morphology of modelled and observed green line emission when using a Lindzen gravity wave scheme in the TIME-GCM. The amplitude of the gravity wave spectrum was however tuned to give good agreement between model winds and the UARS wind data of *McLandress* [1996b]. This would remove the need to increase the lower boundary tidal forcing beyond that suggested by the GSWM. The magnitude of the modelled airglow was however still less than that observed by a factor of 2-3. The

authors suggested that this discrepancy might be related to the gravity wave parameterisation used, or the magnitude of the diurnal tide. To date, GCM studies of $O(^1S)$ emission have not been able to model the absolute value of $O(^1S)$ emission rates observed by WINDII.

It is clear that the representation of gravity waves in global GCMs has a significant impact on modelling studies of the MLT region. In this chapter two different gravity waves parameterisations have been implemented in the CMAT GCM. A hybrid *Lindzen-Matsuno* scheme developed by *Meyer* [1999a,b], and the *Medvedev and Klaassen* [2000] scheme. Comparison is made between the zonally averaged zonal wind in the MLT region produced when using the *Meyer* [1999a,b] scheme and the *Medvedev and Klaassen* [2000] scheme during the northern hemisphere spring equinox. The same vertical profile of eddy diffusion is used with both schemes so as to prevent any indirect influences on the tide that may arise from changes to eddy mixing (*Meyer* [1999a], *Akmaev* [2001]). The meridional winds and tidal amplitudes resulting from the calculated gravity wave momentum deposition from both schemes are assessed. To demonstrate the effect of including a self-consistently calculated eddy diffusion profile, a comparison between zonal and meridional winds calculated using the standard CMAT eddy diffusion profile and one derived from *Medvedev and Klaassen* [2003] is presented. The impact of the gravity wave parameterisations on calculated atomic oxygen distribution and green line emission rates is also discussed. Model results are compared to observations from the HRDI and WINDII instruments on the UARS satellite which measure wind speeds, tidal amplitudes and $O(^1S)$ emission rates. *Burrage et al.* [1996] and *McLandress* [1996a] showed that wind speeds measured by these two instruments were in close agreement.

6.2 THE OBSERVED DIURNAL TIDE

The diurnal tide is a thermally driven westward propagating oscillation with a period of 24 hours. Satellite observations of horizontal winds have measured diurnal tidal wind speeds of up to $70ms^{-1}$ with a vertical wavelength of around 25km (*Hays et al.* [1994], *Burrage et al.* [1996], *McLandress et al.* [1996b]). Maximum horizontal wind amplitudes are observed between 15° and 30° latitude at altitudes close to the

mesopause. Temperature amplitudes in excess of 30K have been measured and occur near the equator (*Hays et al.* [1994]). Above 95km, the tide begins to dissipate and wind speeds decrease. *McLandress* [1998] suggested that dissipative mechanisms in the lower thermosphere may be associated with gravity wave breaking.

In general, peaks in meridional wind amplitudes reported by instruments onboard the UARS satellite (*Hays et al.* [1994], *McLandress et al.* [1996b]) are substantially larger than those reported by MF radar at low to mid latitude (*Fritts and Isler* [1994], *Potnyagin* [1999], *Pancheva et al.* [2002]). *Meyer* [1999a] suggested this discrepancy could be related to the presence of the non-migrating tide, shown to have a suppressing effect in the Pacific sector (*Hagan et al.* [1996]). The cause of this offset is still unclear. Studies by *Harris et al.* [2002] and *Roble and Shepherd* [1997] have shown that the morphology of tidally influenced features such as the $O(^1S)$ green line emission cannot be realistically recreated when tidal amplitudes of the scale measured by MF radar are used. Therefore for the purpose of this study, we shall assume that the winds and amplitudes measured by UARS are correct.

Both ground based and satellite observations of the diurnal tide demonstrate a high degree of temporal and spatial variability in tidal amplitude. A pronounced semiannual amplitude oscillation is present at subtropical latitudes where maxima (minima) are present during equinox (solstice). The phase of the observed tide is seen to be fairly constant throughout the year (*Fritts and Isler* [1994], *Hays et al.* [1994], *McLandress* [1996b]).

6.3 ATOMIC OXYGEN $O(^1S)$ GREEN LINE VOLUME EMISSION

It is generally accepted that atomic oxygen green line emission arises mainly through the following process

$$O + O + M \rightarrow O_2^* + M$$

$$O_2^* + O \rightarrow O(^1S) + O_2$$

$$O(^1S) \rightarrow O(^1D) + h\nu\,(557.7nm)$$

where O refers to the ground state of atomic oxygen ($O(^3P)$). $O(^1D)$ [AD290]and $O(^1S)$ are the first and second excited states respectively. In the three-body recombination reaction, M is usually assumed to be either N_2 or O_2. This mechanism is often referred to as the Barth mechanism (*Barth and Hildebrand* [1961]) and is the mechanism assumed for calculation of $O(^1S)$ emission rates in CMAT. The exact mechanisms involved in the creation and loss of the excited states are still not well known.

McDade et al. [1986] presented a set of empirical parameters from which atomic oxygen numbers densities can be adequately derived from airglow observations by using the following expression

$$V_{5577} = \frac{A_5 k_1 [O]^3 \{[N_2]+[O_2]\}}{\{A_6 + k_5[O_2]\}\{C'^{O_2}[O_2] + C'^{O}[O]\}} \qquad 6.1$$

where A_5 is the green line transition probability, k_1 is the rate coefficient for three-body recombination of atomic oxygen, A_6 is the $O(^1S)$ inverse radiative lifetime and k_5 is the rate coefficient for quenching of $O(^1S)$ by atmospheric O_2. C'^{O_2} and C'^{O} are the empirical $O(^1S)$ excitation parameters (related to the properties of the unidentified precursor to $O(^1S)$). $[X]$ is the number density of species X in cm^{-3}. V_{5577} is the green line volume emission rate in photons $cm^{-3}s^{-1}$. This expression is used to calculate the green line volume emission rate in CMAT, as outlined in *Murtagh et al.* [1990].

The parameters given by *McDade et al.* [1986] are based on observations from the ETON (Energy Transfer in Oxygen Nightglow) rocket flights (*Greer et al.* [1986]), that measured both airglow and atomic oxygen simultaneously. The choice of coefficients is dependent on the profiles of O_2, N_2 and temperature used, which came from various neutral atmosphere models. In CMAT, the coefficients used are those derived from use of the MSIS-83 neutral atmosphere model, as given in Table 6.1.

Rate	Reaction	Coefficient
A_5	$O(^1S) \rightarrow (O^1D) + h\nu$ (557.7nm)	1.18 s^{-1}
A_6	$O(^1S) \rightarrow O(^1D,^3P) + h\nu$	1.35 s^{-1}
k_1	$O + O + M \rightarrow O_2^* + M$	$4.7 \times 10^{-33} (300/T)^2$ $cm^{-6}s^{-1}$
k_5	$O(^1S) + O_2(X^3\Sigma) \rightarrow O + O_2$	$4.0 \times 10^{-12} \exp(-865/T)$ $cm^{-3}s^{-1}$
Excitation Parameter		
C^{1O_2}	15	
C^{1O}	211	

Table 6.1 Rate coefficients and excitation parameters for calculation of $O(^1S)$ green line emission in CMAT. For rate references see *Murtagh et al.* [1990].

6.4 CMAT MODEL CONFIGURATION

For each of the simulations presented in this chapter, the CMAT GCM was run to steady state at northern spring equinox under conditions of low solar and geomagnetic activity (F10.7 = 76, $K_p = 2^+$). The CMAT configurations used in the first three model runs differ only in the representation of momentum deposition due to gravity wave breaking. In the first simulation no gravity wave drag is applied. In the second, momentum terms are derived from the *Matsuno-Lindzen* hybrid gravity wave parameterisation of *Meyer* [1999a,b], as described in section 2.13.3. In the third simulation the gravity wave drag applied is calculated using the scheme of *Medvedev*

and Klaassen [2000], as outlined in section 3.9. In the discussion that follows the two gravity wave schemes will be denoted 'M99' and 'MK2000'.

Several studies have shown that changes to the eddy diffusion profile can have a significant impact on the diurnal tide, turbulent mixing being particularly important for tidal dissipation in the MLT (e.g. *Akmaev* [2001], *Mayr et al.* [2001]). The study of *Meyer* [1999a] is one of the few experiments in which eddy diffusion and gravity wave momentum deposition terms have been isolated and their relative effect on the diurnal tide assessed. A similar approach has been adopted in this study; the aim being to compare the impact of the calculated gravity wave drags in each model simulation, excluding any indirect effects that may arise from changes to the eddy diffusion profile. The global mean height dependent eddy diffusion coefficient of *Roble* [1995] is therefore used in all three model runs described above.

In order to demonstrate the impact of self-consistently calculating the eddy diffusion profile, a fourth model simulation has been carried out that uses an eddy diffusion profile derived using the formulation of *Medvedev and Klaassen* [2003], as described in section 3.9.2. The resultant gravity wave induced eddy diffusion profile has been globally averaged to avoid numerical difficulties that arise when the 3D solution is implemented. For the purpose of this study, the height variation and magnitude of the eddy mixing are considered more important than the horizontal and season variations.

6.5 RESULTS AND DISCUSSION

6.5.1 CMAT MEAN ATMOSPHERIC STRUCTURE

The daily mean zonal wind in the MLT region, as calculated by CMAT for each model simulation and as given by the empirical Horizontal Wind Model (HWM) (*Hedin et al.* [1993]) is shown in Figure 6.1. The mid latitude eastward jets arise as a consequence of the balance between pressure gradient and coriolis forces. It is clear from Figure 6.1 (a) that in the absence of gravity wave drag these jets are not closed off in the mesosphere, but continue undamped into the lower thermosphere where ion drag prevents any further growth. At equatorial latitudes where the effect of the Earth's

rotation is weak, a westward flow develops that is thought to be driven by the dissipating diurnal tide [a293](e.g. *Lindzen* [1981]).

When momentum deposition associated with gravity wave breaking is applied to the model, the velocity of the mid-latitude zonal jets is greatly reduced. Wave drag calculated using the M99 scheme (Figure 6.1 (b)) prevents any increase in the jet velocity at altitudes above around 55km, where the velocity reaches a maximum of approximately 40ms^{-1}. Very little structure is seen between ±30° latitude. The velocity of the mid-latitude jets modelled using momentum deposition from the MK2000 scheme (Figure 6.1 (c)) reach a maximum of 45ms^{-1}, between 60 and 65km altitude. A reversal from eastward to westward occurs at about 85-90km, in agreement with UARS observations presented by *McLandress et al.* [1996a]. The maximum jet velocity given by HWM (Figure 6.1 (d)) is approximately 30ms^{-1}, somewhat lower in magnitude than that calculated using both gravity wave drag schemes. The altitude of the peak calculated using the M99 and MK2000 parameterisations is in reasonable agreement with HWM and the UARS observations of *McLandress et al.* [1996a]. A significant difference between the mean zonal winds at equatorial latitudes is visible in Figure 6.1 (b) and Figure 6.1 (c). The equatorial westward flow modelled using the MK2000 scheme is not present when the M99 scheme is used. This modelled westward flow reaches a maximum of about 25ms^{-1} at around 85km altitude. A similar westward flow is present in the HWM data where a maximum velocity of 20ms^{-1} is seen close to 80km altitude.

The zonal mean neutral temperature calculated by CMAT using the M99 and MK2000 schemes is shown in Figure 6.2 along with an MSIS temperature profile generated under similar condition. There is good agreement between the MSIS temperatures and those calculated by CMAT when using both the M99 and MK2000 schemes. An increase in wavelike structures is apparent in the mid- to low-latitude mesopause region when using the MK2000 simulation. This is most likely related to deposition of gravity wave momentum at low latitudes, induced by the MK2000 scheme but not the M99 scheme which calculates almost zero drag between ±20°.

The meridional wind component at 12:00LT (midday) is shown in Figure 6.3 for CMAT simulations utilizing the M99 and MK2000 gravity wave parameterisations. Clearly visible is the propagating diurnal tide with its asymmetric cell-like structure between ±40°. The vertical wavelength of the simulated diurnal tide is approximately 30km when using M99, but only 20 to 22km when using the MK2000 scheme. Observations of the diurnal tidal such as those from HRDI (Figure 6.4) show a vertical wavelength between 20 and 25km. The MK2000 scheme clearly acts to reduce the wavelength of the diurnal tide in the CMAT model. A gravity wave induced reduction in the vertical wavelength has also been report by *Mayr et al.* [2001] who presented a study using the *Hines* [1997a] Doppler spread gravity wave parameterisation. As noted by *Harris et al.* [2002], the HRDI tidal amplitudes of up to 80ms^{-1} are much greater than those modelled by CMAT when the M99 gravity wave scheme is used. Maximum tidal amplitudes modelled using the MK2000 scheme are between 40 and 80ms^{-1}, in good agreement with HRDI observations.

6.5.2 EFFECT OF PARAMETERISED GRAVITY WAVE DRAG ON THE DIURNAL TIDE

In Figure 6.5 the amplitude of the diurnal tide at 12:00LT is shown, as calculated by the CMAT simulations without gravity wave drag (*top*), with drag from the M99 gravity wave scheme (*middle*) and using drags from the MK2000 scheme (*bottom*). In the absence of gravity wave drag, a maximum in diurnal amplitude of about 70ms^{-1} occurs at around 115km altitude, between ±20° and ±30° latitude. When gravity wave forcing from the M99 scheme is included in CMAT, the maximum diurnal amplitude is reduced to about 25ms^{-1} and the location of the peak is shifted to an altitude of around 112km. A weaker secondary peak is visible at about 85km. Diurnal amplitudes modelled using the MK2000 scheme peak at nearly 80 ms^{-1} at an altitude of 110km. Again, a secondary peak is seen, this time centred on 95km. WINDII observations of diurnal tidal amplitude from combined meridional wind data spanning March/April 1992/93 are shown in Figure 6.6 (from *Akmaev* [1997]). A reasonable agreement between observed and modelled amplitudes is seen when momentum terms from the MK2000 scheme are used in CMAT, although the magnitude of the maximum at 110km is slightly higher that that observed. In the absence of gravity wave drag, the

locations of the maxima are too high, whereas with the M99 scheme the magnitude of the tide is a factor of 2 to 3 too low. In all model runs, the magnitude of the secondary maximum at around 95km is lower than that observed. Changes in the altitudes of the amplitude peaks in each model run arise as a result of changes to the vertical wavelength, driven by gravity wave breaking (*Mayr et al.* [2001]).

It is clear from Figure 6.5 that the inclusion of M99 gravity wave momentum terms in CMAT leads to a reduction in the amplitude of the diurnal tide. This was first reported by *Harris et al.* [2002], and supported the results of *Meyer* [1999a] who observed a reduction in tidal amplitude when using the scheme in a 2-D tidal model[a294]. This can be explained by looking at the phase of the gravity wave forcing relative to the phase of the tidal wind field. If gravity wave drag and the tidal wind are more in-phase than out-of-phase, the forcing will result in an acceleration of the wind. Parameterised gravity waves dissipating close to and below the wind maxima [a295](in the direction of gravity wave propagation) will therefore result in an increase in tidal amplitude and an induced downward phase motion (*Fritts* [2003]). If however gravity wave forcing and the tidal wind are more out-of-phase than in-phase, dissipation of the gravity wave will result in a reduction in amplitude (*McLandress* [1998]). In this case, the induced phase motion will depend on where the forcing occurs relative to the wind maxima. The phase of the meridional winds and gravity wave forcing at 20°S as calculated by CMAT using the M99 and MK2000 schemes is shown in Figure 6.7. Gravity wave drag calculated by the MK2000 scheme is nearly in-phase with the CMAT meridional winds and thus causes an acceleration of the background wind. *Manson et al.* [2002] also reported an increase in tidal amplitude when using the MK2000 scheme in the CMAM GCM (*Beagley et al.* [1997]). The nearly out-of-phase relationship between meridional winds and gravity wave forcing calculated by the M99 scheme leads to a reduction in the tidal amplitude, as reported by *Meyer* [1999a] and *Harris et al.* [2002].

Figure 6.8 shows the location of the diurnal tide maxima at 20°S as a function of local time for each CMAT simulation. The introduction of gravity wave forcing results in an advancement[AD296] of the tide such that at a given height the peak amplitude occurs at an earlier time of day. The induced phase motion depends on where gravity wave

forcing occurs relative to the wind maxima. The nearly out-of-phase relationship between the meridional winds and gravity wave drag calculated using the M99 scheme illustrated in Figure 6.7 results in the greatest advancement of the diurnal tide at low altitudes. At altitudes between 105 and 120km where the diurnal tide peaks, both gravity wave schemes cause an advancement of the tidal phase of between 6 and 7 hours. Similar phase shifts have also been modelled by *Meyer* [1999a], *McLandress* [1998], and *Mayr et al.* [2001] when using a variety of gravity wave parameterisations.

6.5.3 IMPACT OF SELF-CONSISTENTLY CALCULATED EDDY DIFFUSION ON ZONAL AND MERIDIONAL WINDS

In Figure 6.9, the global mean eddy diffusion profile used in the first 3 CMAT model simulations is shown (*top*), along with that calculated using the expression of *Medvedev and Klaassen* [2003] (*middle*), henceforth denoted as MK2003. The global mean eddy diffusion calculated using the *Hines* [1997a] gravity wave scheme is also shown (from *Akmaev* [2001]) (*bottom*). Both the standard CMAT profile due to *Roble* [1995] and the MK2003 profile peak at around 105km altitude, whereas the *Hines* equinox profile peaks closer to 110km. The standard CMAT profile has a maximum value of about 70 m^2s^{-1}, slightly higher than the 58 m^2s^{-1} predicted using the MK2003 formulation. Both of these values are notably lower that that derived from the *Hines* scheme which predicts a maximum of just over 100 m^2s^{-1}. All the profiles drop off rapidly above the peak altitude where molecular dissipation starts to dominate. Below the peak the rapid drop in eddy diffusion seen in all profiles is most pronounced in the MK2003 global mean where a value of less than 0.1 m^2s^{-1} is predicted at 60km altitude. The standard CMAT profile however never drops below about 4 m^2s^{-1} and contains a small peak at around 50km altitude that is not present in the self-consistently calculated profiles. There are clearly significant differences between the two eddy diffusion profiles used in CMAT, and between the global mean profiles calculated using the *Hines* and MK2003 formulations.

The impact of using the MK2003 global mean eddy diffusion coefficient in CMAT can be seen by comparing the CMAT zonal and meridional winds calculated using this profile, with CMAT winds calculated using the standard *Roble* [1995] profile. Figure

6.10 shows the diurnal mean zonal wind, the meridional wind at 12:00LT, 0° longitude, and the diurnal tidal amplitude at 12:00LT predicted by CMAT when using a global mean of the MK2003 eddy diffusion profile, and gravity wave drag terms from the MK2000 scheme. The diurnal mean zonal wind shows little deviation from that calculated using the standard eddy diffusion profile (shown in Figure 6.1 (c)), both displaying maximum mid-latitude zonal jet velocities of around 45 ms^{-1} between 60 and 65km altitude. The equatorial westward flow is slightly stronger when using the self-consistently calculated eddy diffusion and extends to approximately 100km altitude, about 5 km higher than when the standard profile is used. The greatest differences are seen above 100km where the mid-latitude westward flows are stronger in both the northern and southern hemispheres when using the MK2003 global mean profile.

A comparison of the meridional winds at 12:00LT, 0° longitude, as shown in Figure 6.3 (*bottom plot*) and Figure 6.10 (*middle plot*) shows that there is little difference between the winds calculated using the CMAT standard eddy diffusion profile and that derived using the expression of MK2003. A small increase in the magnitude of the meridional winds is visible at mid-latitudes. This can be more easily assessed by looking at the diurnal tidal amplitude, shown in Figure 6.10 (*bottom plot*). Maximum wind speeds of 80 ms^{-1} occur between approximately 105 and 112km altitude in both hemispheres. A secondary maximum of 80ms^{-1} is also visible between 97 and 100km in the southern hemisphere. While the locations of the maxima are approximately the same as in Figure 6.5 (*bottom plot*), the magnitudes are greater when using the MK2003 eddy diffusion profile. This is due to the low eddy diffusion coefficient produced by the self-consistent calculation, as discussed above. A reduction in the eddy mixing at altitudes where the amplitude of the diurnal tide peaks results in less tidal dissipation and thus stronger tidal winds. As previously mentioned, the turbulent mixing applied in CMAT via the *Roble* [1995] or MK2003 eddy diffusion profiles is lower than that derived using the gravity wave scheme of *Hines* [1997a]. These larger eddy mixing coefficients have been used in the studies of *Akmaev* [2001] and *Mayr et al.* [2001]. Inclusion of a similar eddy diffusion coefficient in CMAT is expected to result in a reduction of the diurnal tidal amplitude.

6.5.4 EFFECT OF PARAMETERISED GRAVITY WAVE DRAG ON EQUATORIAL ATOMIC OXYGEN GREEN LINE EMISSION RATES

Observations such as those made by the HRDI and WINDII instruments onboard the UARS satellite have revealed a large degree of spatial and temporal variability in the atomic oxygen green line volume emission rate. Clearly visible in nightglow observations of the MLT region is a strong local time dependence, associated with the migrating diurnal tide (e.g. *Yee et al.* [1997], *Shepherd et al.* [1995]). This tidal signature is thought to be primarily the result of variations in the atomic oxygen mixing ratio, driven by tidally induced vertical winds (*Ward* [1998], *Ward* [1999]). The representation of the diurnal tide in CMAT is strongly affected by the gravity wave parameterisation used, as discussed in the previous section. The atomic oxygen concentration and associated green line emission will therefore also be affected by the way gravity wave momentum deposition is calculated in the model.

The local time variation of green line volume emission between 18:00LT and 6:00LT at the equator, as calculated in simulations 2 and 3 is shown in Figure 6.11. The model runs used the M99 and MK2000 gravity wave schemes respectively, with the standard CMAT eddy diffusion profile. The volume emission rate measured by WINDII in March/April 1993 is also shown (from *Shepherd et al.* [1995]). The WINDII data reveals a bright airglow layer in the evening with a peak of about 200 photons $cm^{-3}s^{-1}$, starting at around 95km and descending with time to around 89km. From 20:00LT to midnight, the peak weakens and moves up in height. Past 2:00LT the altitude of the peak remains steady at about 96km and its strength increases. By 6:00LT the magnitude of the peak has reached about 200 photons $cm^{-3}s^{-1}$. When using gravity wave forcing terms from the M99 gravity wave scheme in CMAT (Figure 6.11 (*top*)), a peak in green line emission of about 160 photons $cm^{-3}s^{-1}$ is present in the evening sector at a height of about 98km. The altitude of the peak is fairly constant throughout the 12-hour period, varying by just a few km. The magnitude of the peak decreases throughout the evening, dropping to a minimum of approximately 120 photons $cm^{-3}s^{-1}$ at 22:00LT. Very little tidal structure is visible due to the damping effect of the gravity wave parameterisation on the diurnal tidal wind. The evening sector peak in green line emission modelled using the MK2000 scheme in CMAT (Figure 6.11 (*middle*)) occurs

at around 94km and reaches nearly 220 photons $cm^{-3}s^{-1}$. The temporal variation of the airglow layer displays a distinct tidal signature, resulting in better agreement with WINDII observations than when the M99 scheme is used. Minimum values of around 50-60 photons $cm^{-3}s^{-1}$ present in the early morning sector of the WINDII observations are also seen in the MK2000 simulation results. A downward progression of the peak with time is present in the MK2000 simulation, although the altitude of the peak does not extend below 90km[AD298]. The phase of the MK2000 simulated nightglow lags behind that of the WINDII observations by approximately 4-5 hours. This can be explained by the tidal phase shift induced by the gravity wave scheme as illustrated in Figure 6.8.

When using the MK2000 gravity wave scheme in conjunction with a global mean eddy diffusion profile derived using the expression of MK2003, peak green line volume emission rates are greater than those modelled using the standard CMAT eddy diffusion profile (not shown). Maxima of over 280 photons $cm^{-3}s^{-1}$ are predicted to occur in the evening sector. The diurnal variation of the airglow however does not differ significantly from that modelled with the *Roble* [1995] eddy diffusion profile. The increase in peak volume emission rate is most likely caused by an excess abundance of atomic oxygen rich air, being transported down from the thermosphere by tidally induced vertical winds (*Ward* [1998], *Ward* [1999]). The increased tidal amplitude resulting from reduced eddy induced dissipation is expected to cause an increase in the advective vertical winds. Both of the global mean eddy diffusion profiles used in CMAT have peak magnitudes less than those used in the studies of *Akmaev* [2001] and *Mayr et al.* [2001], derived from the gravity wave parameterisation of *Hines* [1997a]. An increase in the turbulent dissipation rates used in CMAT in line with those used by *Akmaev* [2001] and *Mayr et al.* [2001] would result in a decrease in the diurnal tidal amplitude. Under such conditions, an associated decrease in the peak green line volume emission rate would be anticipated.

A comparison of local time variation in atomic oxygen concentration calculated by CMAT using the M99 and MK2000 schemes is shown in Figure 6.12. Using the M99 scheme, CMAT predicts a maximum number density of around $5x10^{17}$ m^{-3} at an

altitude of about 98km. Neither the magnitude of the peak nor the altitude at which is occurs show a significant variation throughout the day. The CMAT simulation using the MK2000 scheme produces a peak atomic oxygen number density of nearly $7x10^{17}$ m^{-3} which is seen to descend with time from 98km at local noon to approximately 90km near midnight. Another layer of atomic oxygen forms near 105km just after midnight, which gradually descends in altitude. A clear tidal structure is present, whereby atomic oxygen is effectively pushed downward from higher altitudes to the upper mesosphere where recombination occurs (*Shepherd* [1995], *Yee*[a299] *et al.* [1997], *Ward*[AD300] [1998]). As previously mentioned, tidally induced vertical advection is thought to be the dominant process causing this structure[AD301]. A small increase in vertical winds is present when the MK2000 scheme is used (not shown). This may account for the increased peak atomic oxygen density as more O rich air is being pushed down from the thermosphere. An increase in the equatorial atomic oxygen local maximum was also report by *Harris*[a302] *et al.* [2002] in association with pronounced tidal forcing of atomic oxygen.

While previous modelling studies have been successful in reproducing the observed tidal structure of the green line emission rate, the absolute magnitude of the peak is generally a factor of 2 to 3 less than that observed (*Yee et al.* [1997], *Harris et al.* [2002], *Roble and Shepherd*[AD303] [1997], *Zhang* [2001]). This has been attributed to overestimated eddy diffusion and downward mixing[AD304], underestimated[AD305] tidal wind amplitude and cooling effects [AD306]induced by the dissipation of gravity waves. While it is clear that the choice of gravity wave parameterisation in GCMs can lead to very different airglow profiles, both of the schemes used in CMAT produce green line emission profiles with peak magnitudes that are in reasonable [AD307]agreement with WINDII observations. This is attributed to the improved representation of photochemical processes, thermodynamics and composition in the model, achieved through use of high resolution solar flux data combined with updated photoionisation and absorption cross sections (as discussed in section 3.5.5.3) All of these factors are expected to have had a significant impact on the representation of green line emission in the CMAT model.

6.6 CONCLUSIONS

The CMAT GCM has been used to compare the effect of gravity wave forcing, calculated using two different gravity wave parameterisations, on the mean zonal wind, the diurnal tidal wind, atomic oxygen concentrations and associated O(^{1}S) green line volume emission rates in the MLT region. Three model simulations were performed using a 1D eddy diffusion profile due to *Roble* [1995], the first with no gravity wave momentum deposition, the second using gravity wave forcing from the hybrid *Matsuno-Lindzen* gravity wave scheme of *Meyer* [1999a,b] (M99) and the third using forcing from the scheme of *Medvedev and Klaassen* [2000] (MK2000). All model runs were for March equinox under conditions of low solar and geomagnetic activity.

The magnitude and morphology of the daily mean zonal winds produced by CMAT using the M99 and MK2000 gravity wave schemes are in reasonable agreement with those given by the HWM. Closure of the mid-latitude zonal jets at mesospheric heights is achieved with both schemes. The low latitude westward flow, seen in both UARS and HWM zonal mean wind data, is recreated by CMAT when using the MK2000 scheme but not when the M99 scheme is used.

Meridional winds modelled by CMAT are in reasonable agreement with observations from UARS when gravity wave forcing from the MK2000 scheme is used. The meridional wind is dominated by the propagating diurnal tide and as such can be used as a clear indicator of tidal / gravity wave interactions in the model. As suggested by *Harris et al.* [2002], momentum deposition from the M99 scheme causes a dampening of the diurnal tide in CMAT. As a consequence, peak meridional winds modelled with M99 gravity wave forcing are a factor of 2-3 weaker than those measured by the HRDI and WINDII instruments on the UARS satellite. Comparison of the phase of imposed gravity wave drag and the tidal wind reveals that drag terms from the M99 scheme are nearly out-of-phase with the wind, whereas drag terms from the MK2000 scheme are nearly in-phase with the wind. This relationship is thought to produce the marked reduction in tidal amplitude observed when the M99 scheme is used, and the slight enhancement when the MK2000 scheme is used. The introduction of gravity wave

forcing from both schemes acts to advance the phase of the tide, meaning the maximum wind occurs at an earlier time of day at a given height.

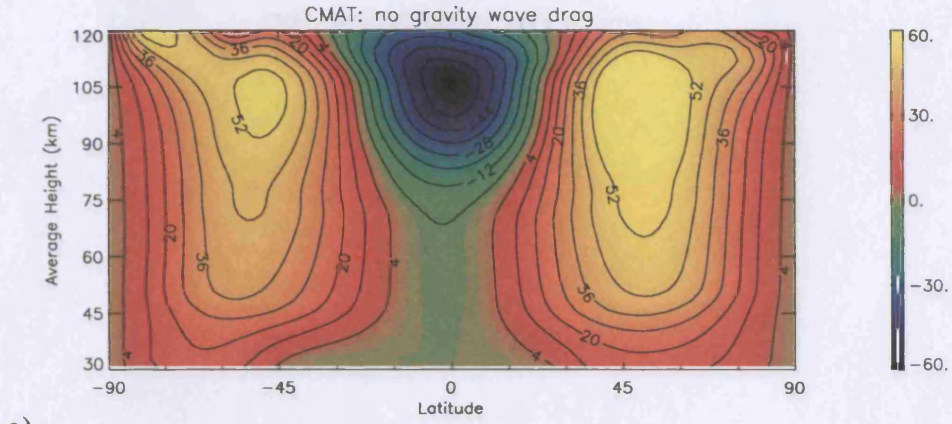
Comparison of the *Roble* [1995] 1D eddy diffusion profile usually used in CMAT with a global mean derived using the expression of *Medvedev and Klaassen* [2003] (MK2003) shows that the self-consistently calculated profile is lower in magnitude than that usually used. Both of these profiles have peak values less than that predicted by the *Hines* [1997a] scheme, used in the studies of *Mayr et al.* [2001] and *Akmaev* [2001]. The diurnally averaged zonal wind calculated by CMAT using the MK2000 gravity wave scheme and the MK2003 eddy diffusion profile does not differ significantly from that calculated when using the 1D eddy diffusion profile of *Roble* [1995]. Application of self-consistently calculated eddy mixing in CMAT results in a small increase in meridional winds, associated with an increase in the diurnal tidal amplitude. This is thought to be the result of decreased tidal dissipation, arising from the reduction in turbulent mixing. The studies of *Meyer* [1999a], *Akmaev* [2001] and *Mayr et al.* [2001] also find a correlation between eddy diffusion and the amplitude of the diurnal tide.

The magnitude and tidal structuring of atomic oxygen 557.7nm volume emission rates modelled by CMAT using the MK2000 gravity wave scheme agree well with WINDII observations. Tidal oscillations penetrate into the peak atomic oxygen layer and lead to a strong diurnal variation in O concentration and green line emission. This is thought to be primarily due to tidally induced changes in vertical advective winds (*Ward* [1998], *Ward* [1999]). The tidal signature is not present in the CMAT nightglow when the M99 scheme is used due to the tendency of the scheme to damp the amplitude of the tide. Good agreement between the absolute magnitudes of modelled airglow and WINDII observations is attributed to the updated representation of solar flux and associated photochemical parameters in the CMAT GCM.

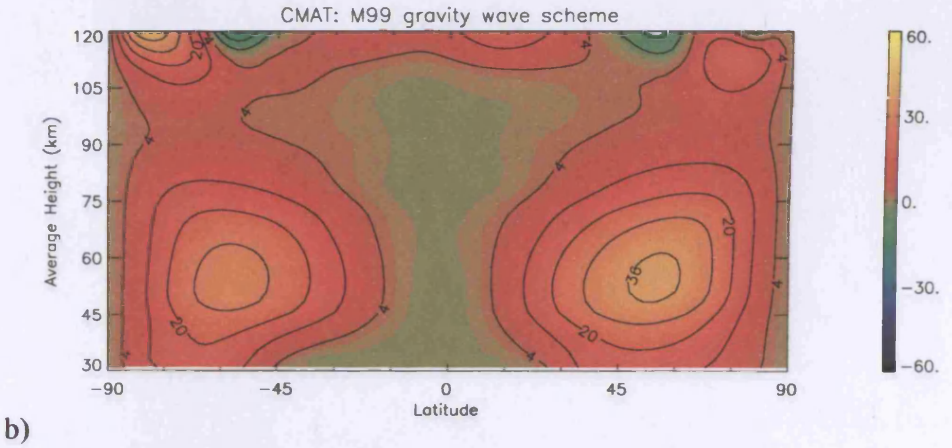
Previous modelling studies have shown that the magnitude of the diurnal tide is strongly influenced by the global eddy diffusion field, and have suggested that momentum deposition by gravity wave breaking also has some influence (*Meyer*

[1999a], *Akmaev* [2001]). In this study, the role of parameterised gravity wave momentum sources has been investigated and they are found to have a significant impact on tidal amplitudes and phase, even when the eddy diffusion profile is held constant. The density of long-lived atomic oxygen in the MLT region is strongly affected by dynamics associated with the diurnal tide. WINDII observed green line emission is therefore also dependent on the tidal winds. The diurnal variability of atomic oxygen green line emission modelled by CMAT is improved by the introduction of gravity wave forcing from the MK2000 scheme.

The development of a time-efficient parameterisation to realistically represent the effects of gravity wave breaking in GCMs is still the subject of much research. 'Tuning' of gravity wave schemes in GCMs to produce reasonable wind profiles is still common, although more robust and complex schemes are being developed that include anisotropic, orographic gravity wave sources and seasonal variation in the gravity wave induced eddy diffusion profile. For simplicity, spatial and temporal variations in gravity wave forcing have not been considered in this study. Inclusion of a realistic gravity wave spectrum and a study of its effect on annual timescales will be a future improvement to the representation of wave-wave interactions in the CMAT GCM.

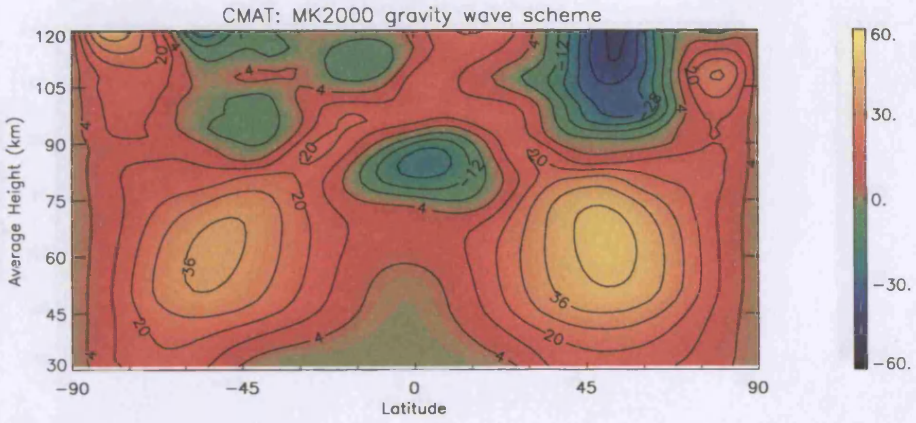


a)

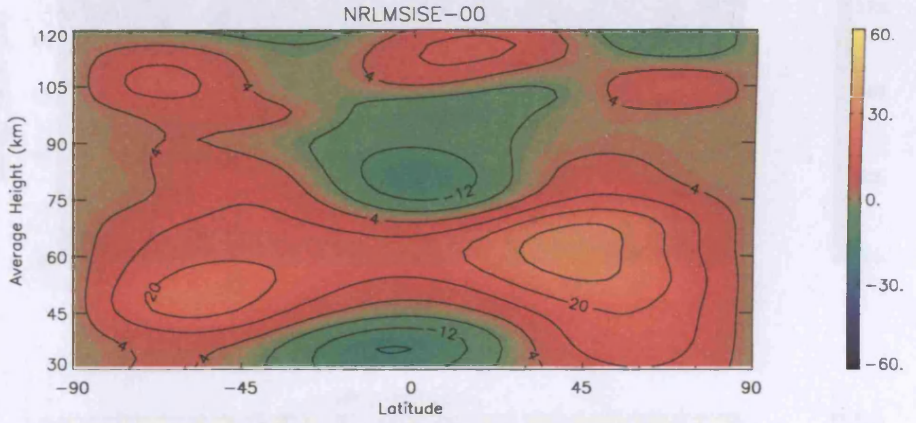


b)

Figure 6.1 Daily mean zonal wind in ms^{-1} as calculated by CMAT with (a) no gravity wave drag (*top*), and (b) gravity wave drag calculated using the MK(99) scheme (*bottom*). Positive values denote eastwards winds. Spring equinox, F10.7 = 76.



c)



d)

Figure 6.1 Daily mean zonal wind in ms^{-1} as calculated by CMAT with (c) gravity wave drag calculated by the MK2000 scheme (*top*) and (d) as given by HWM (*bottom*). Positive values denote eastwards winds. Spring equinox, F10.7 = 76.

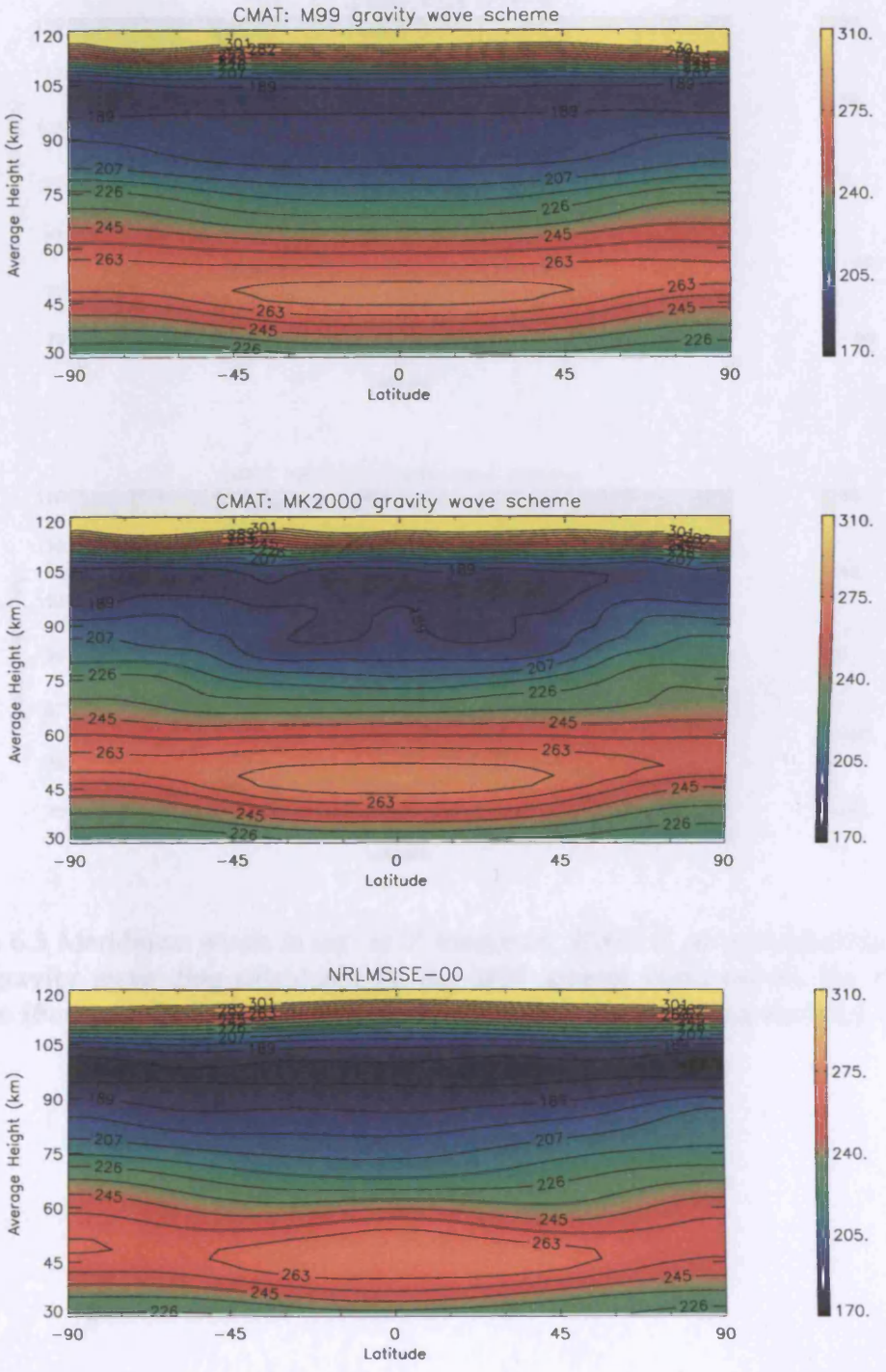


Figure 6.2 Daily mean neutral temperature (K) as calculated by CMAT with gravity wave drag calculated by the M99 scheme (*top*), drag from the MK2000 scheme (*middle*) and as given by MSIS (*bottom*). Spring equinox, F10.7 = 76.

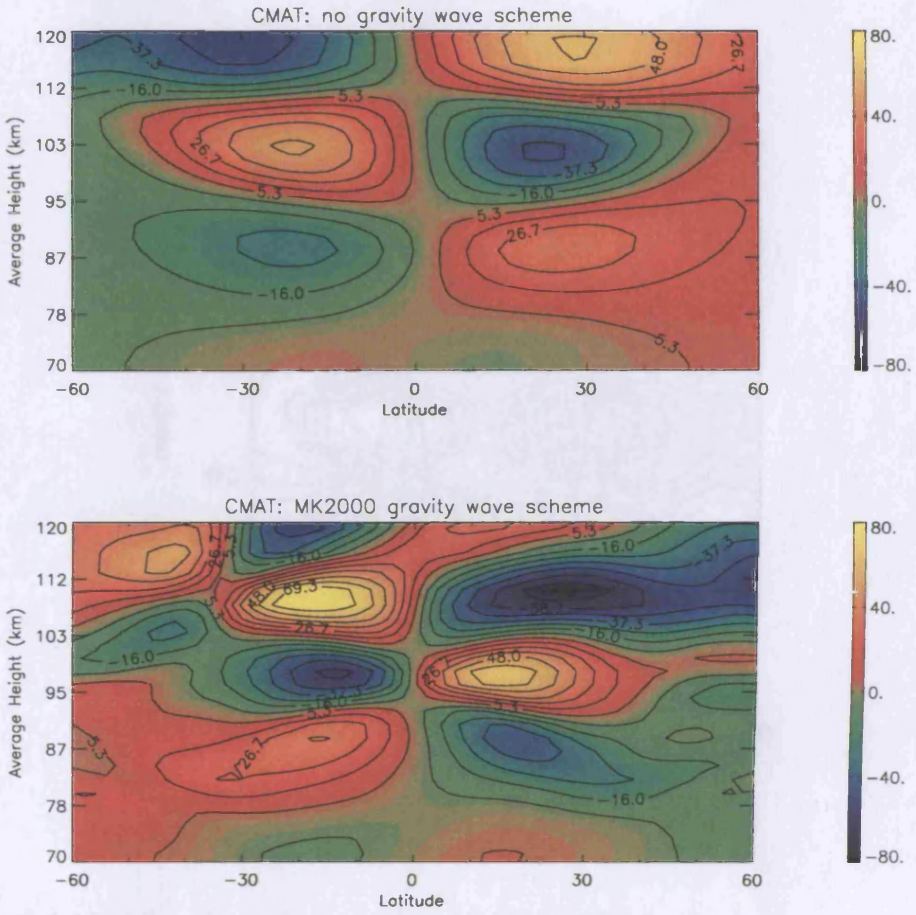


Figure 6.3 Meridional winds in ms^{-1} at 0° longitude, 12:00LT, as calculated by CMAT with gravity wave drag calculated by the M99 scheme (*top*) and by the MK2000 scheme (*bottom*). Positive values denote southward winds. Spring equinox, F10.7 = 76.

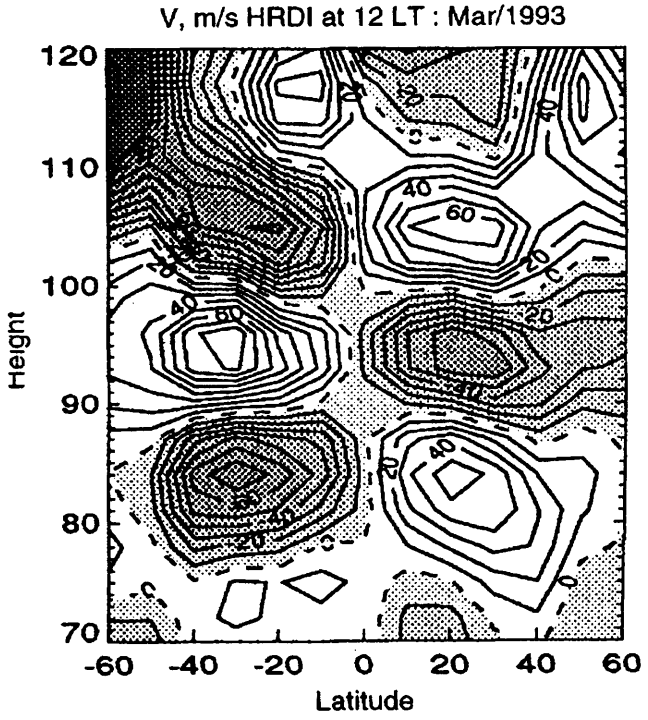


Figure 6.4 Meridional winds in ms^{-1} at 12:00LT for March, as measured by HRDI (from *Yudin et al.* [1997]), Negative values denote southward winds.

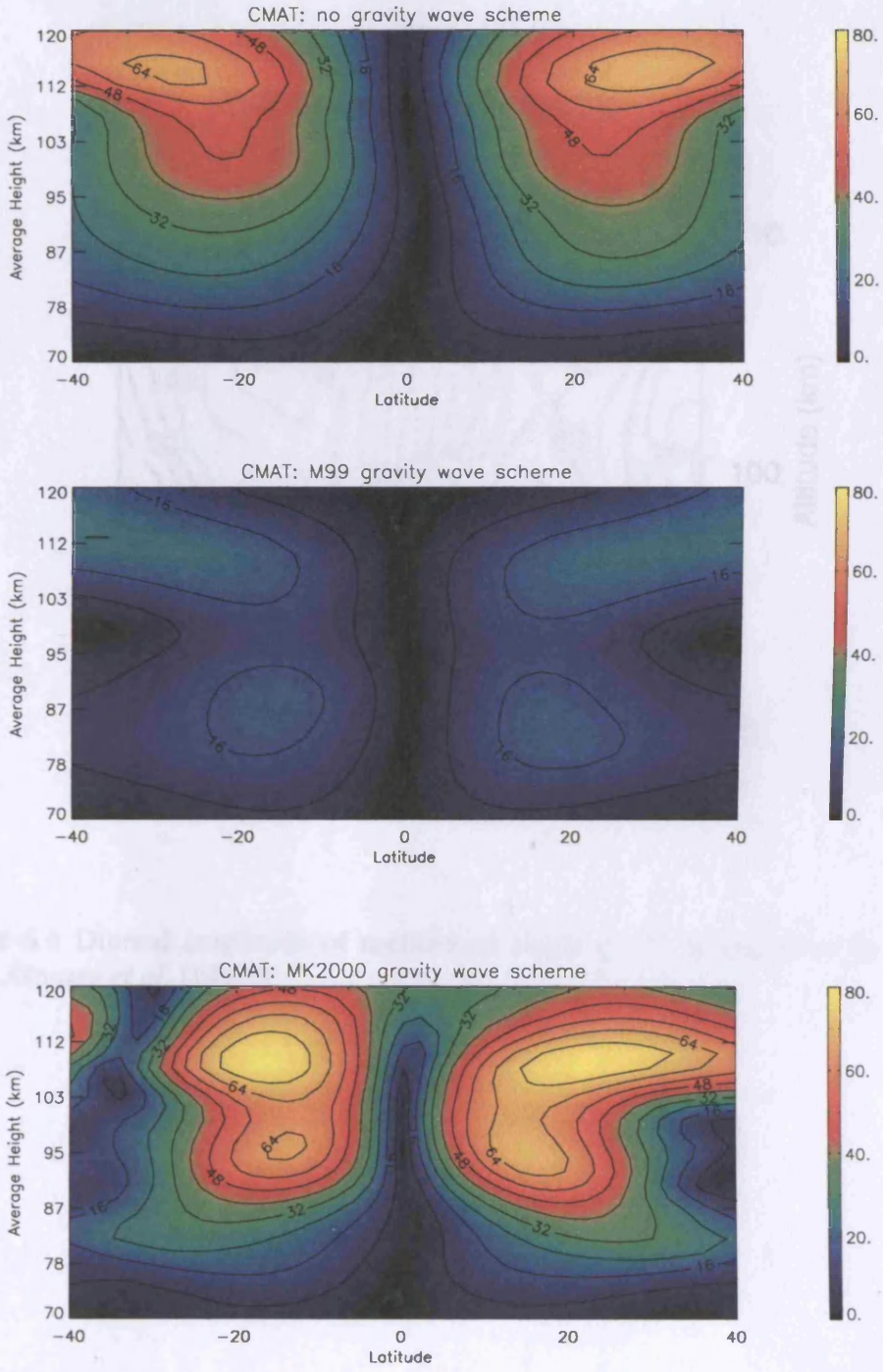


Figure 6.5 Diurnal amplitude of meridional wind (ms^{-1}) at 12:00LT calculated by CMAT with no gravity wave drag (*top*), gravity wave drag from the M99 (*middle*) and MK2000 (*bottom*) schemes. Spring equinox, F10.7 = 76.

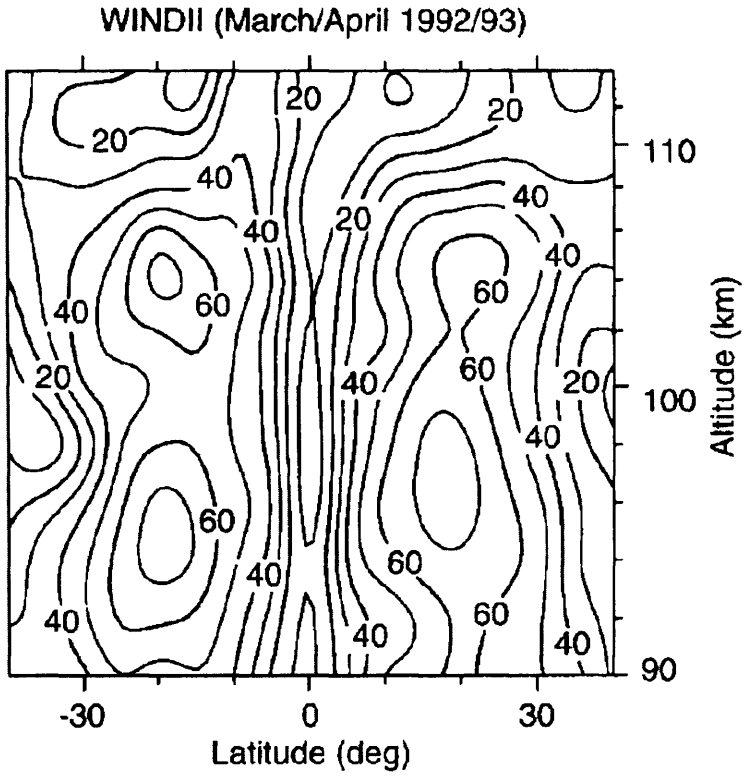


Figure 6.6 Diurnal amplitude of meridional winds (ms^{-1}) as measured by WINDII (*from Akmaev et al.* [1997]).

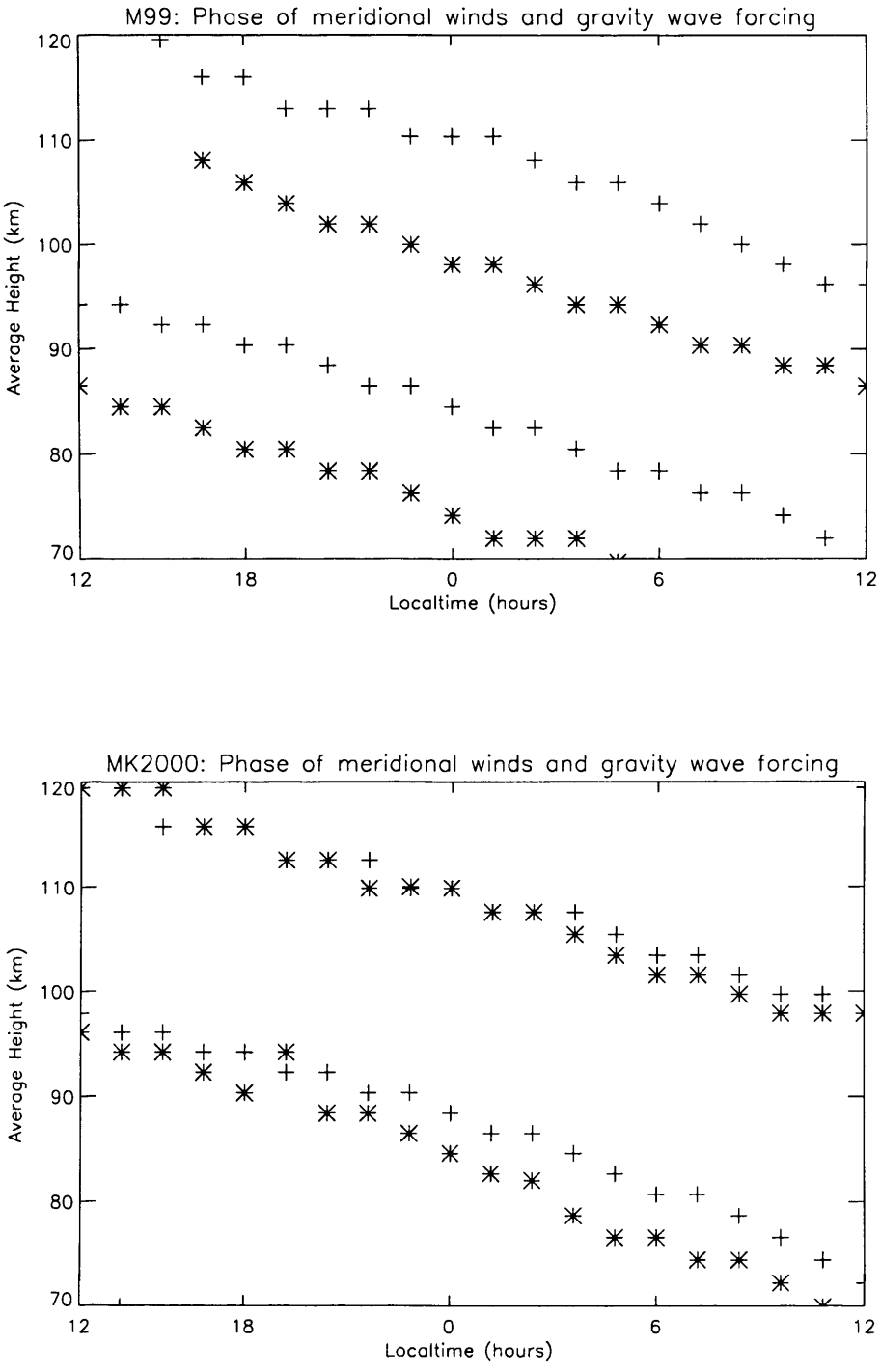


Figure 6.7 Phase (local time of maxima) versus height of meridional winds (*crosses*) and meridional gravity wave forcing (*stars*) at 20°S as calculated by CMAT using the M99 (*top*) and MK2000 (*bottom*) gravity wave schemes. Spring equinox, F10.7 = 76.

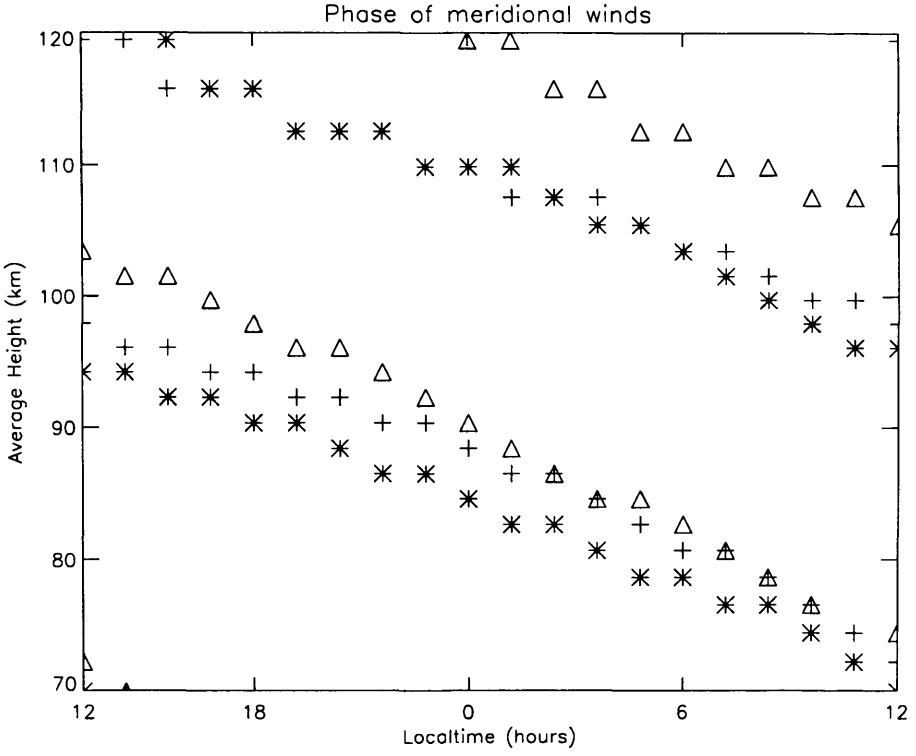


Figure 6.8 Phase of the meridional winds at 20°S as calculated by CMAT with no gravity wave forcing (*triangles*), with gravity wave forcing from the M99 (*stars*) and MK2000 (*crosses*) schemes. Spring equinox, F10.7 = 76.

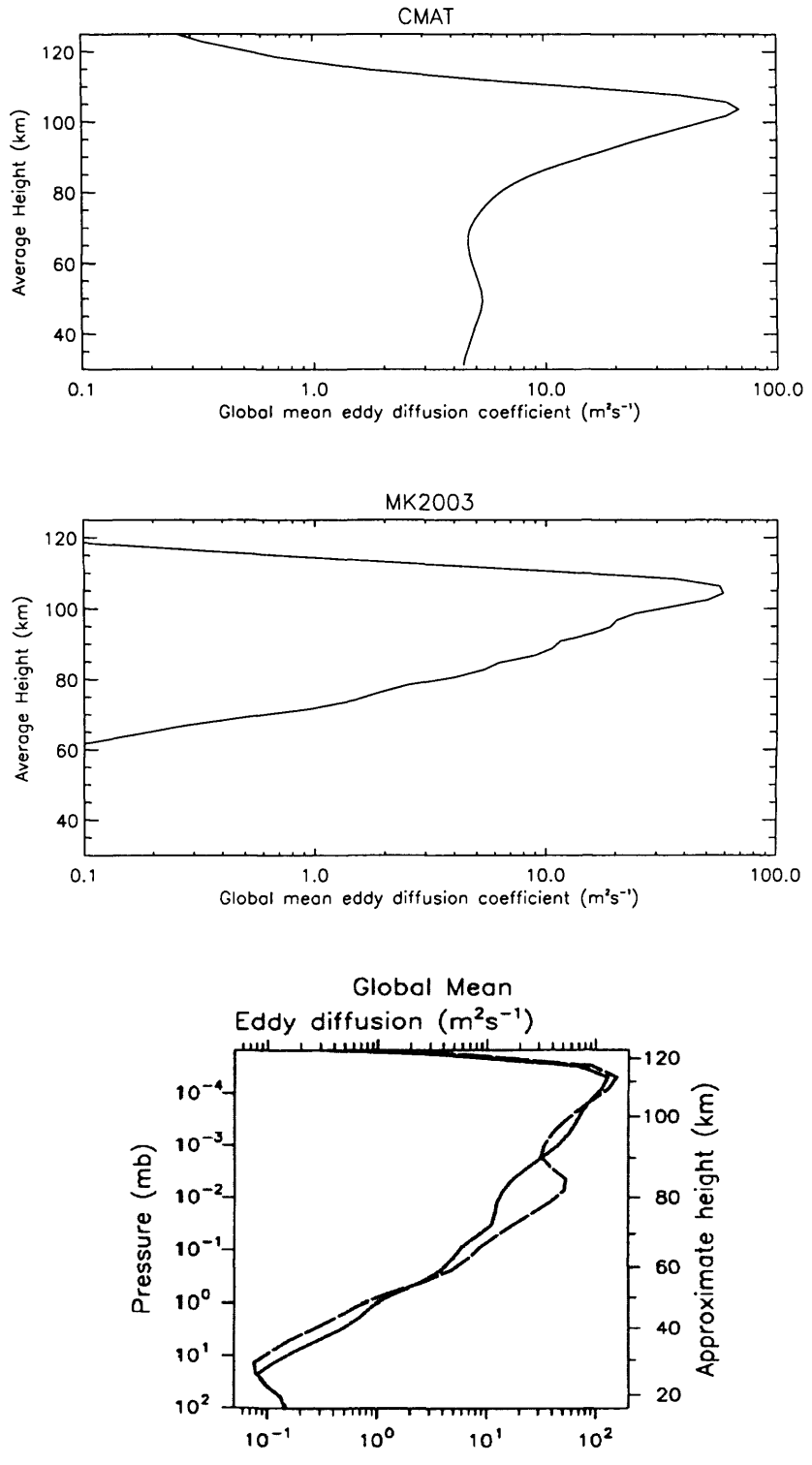


Figure 6.9 From top to bottom, plots of the global mean eddy diffusion profile (m^2s^{-1}) used in CMAT (due to *Roble* [1995]), as calculated using the expression of *Medvedev and Klaassen* [2003], and as calculated using the gravity wave parameterisation of *Hines* [1997a] (from *Akmaev* [2001]) for January (*dashed*) and April (*solid*).

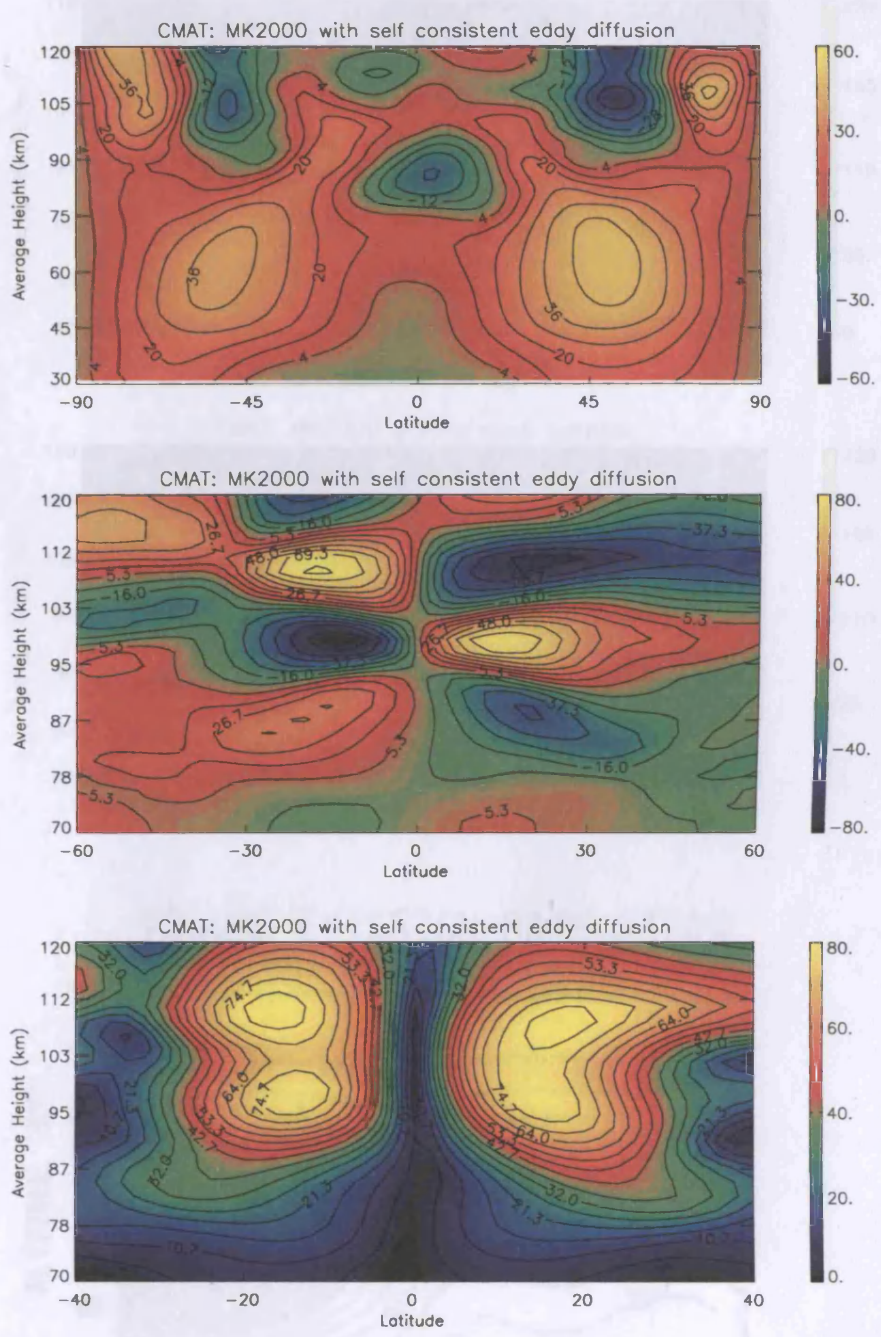


Figure 6.10 CMAT calculated daily mean zonal wind (*top*), meridional winds at 12:00LT (*middle*) and diurnal amplitude of the meridional wind at 12:00LT (*bottom*), using a global mean eddy diffusion profile calculated using the expression of *Medvedev and Klaassen* [2003]. Conditions are appropriate to spring equinox, F10.7 = 76. Positive values denote eastward and southward winds in units of ms^{-1}.

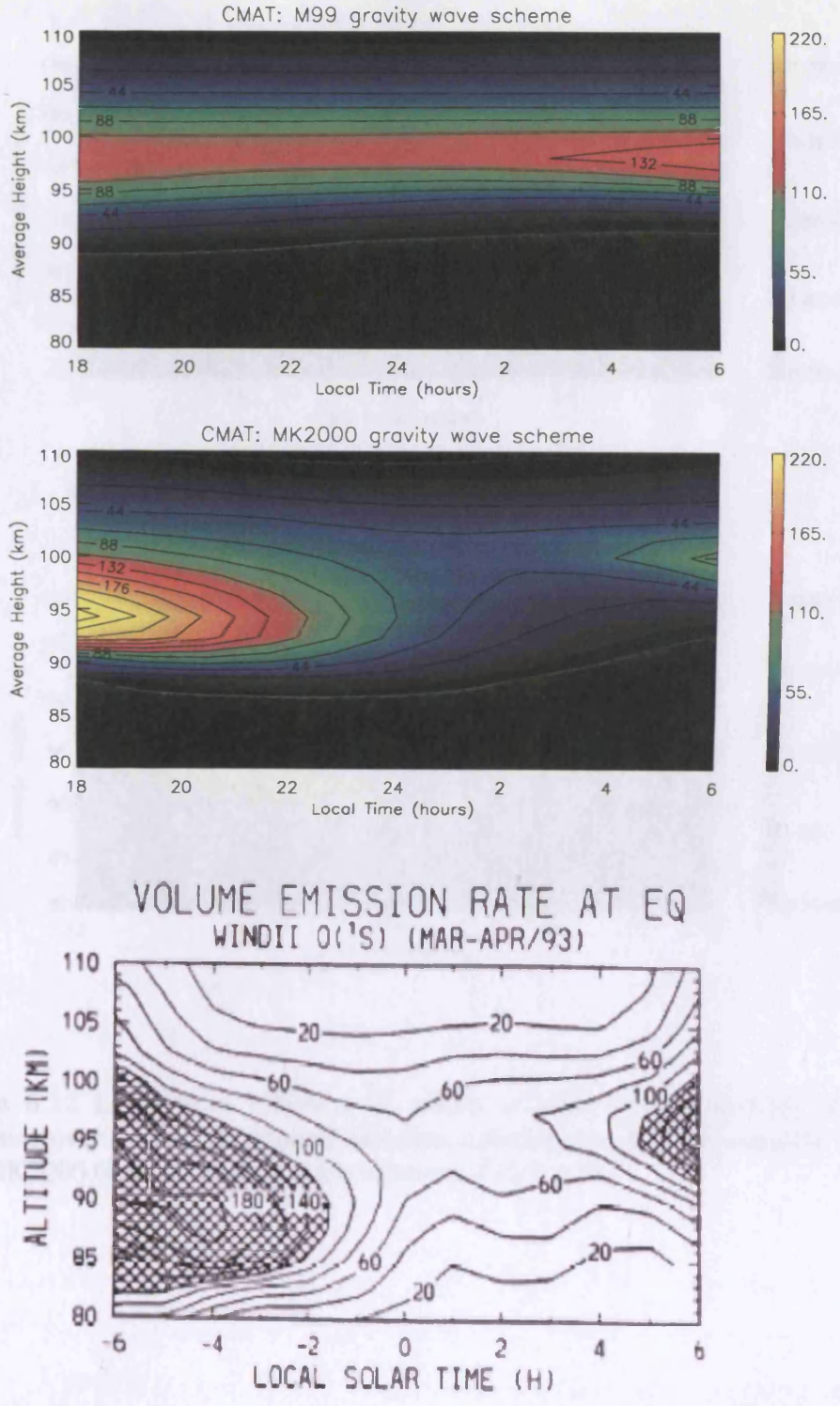


Figure 6.11 Local time variation of 557.7nm green line volume emission rate (photons $cm^{-3}s^{-1}$) at the equator for March equinox: as calculated by CMAT using the M99 scheme (*top*), by CMAT using the MK2000 scheme (*middle*), and as given by WINDII for March/April 1993 (*bottom*) (from *Shepherd et al.* [1995]).

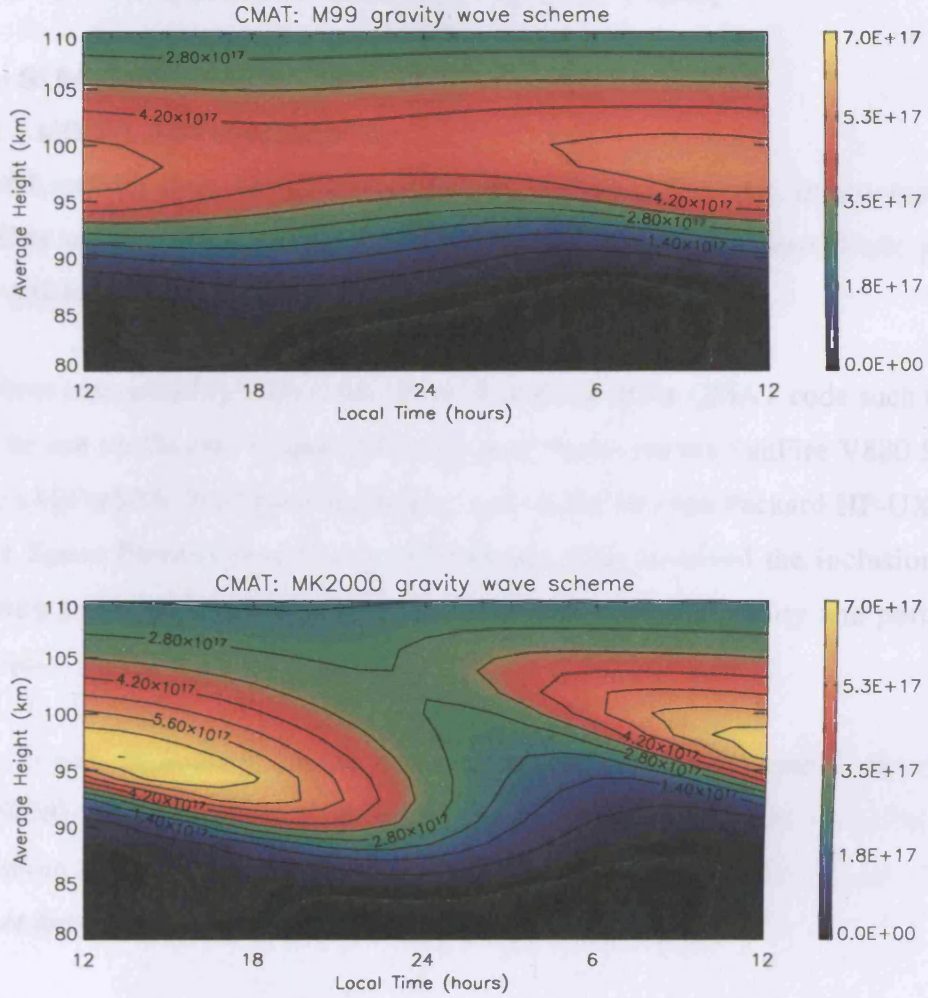


Figure 6.12 Local time variation of atomic oxygen number density (m^{-3}) at 0° longitude over a whole day during March as calculated by CMAT using the M99 (*top*) and MK2000 (*bottom*) gravity wave schemes. F10.7 = 76.

CHAPTER VII. CONCLUSIONS AND FUTURE WORK

7.1 SUMMARY OF WORK

7.1.1 MODEL IMPROVEMENTS

Several updates have been made to the original CMAT model, improving both the usability of the model and the representation of certain key atmospheric processes. Key updates are listed below

Platform compatibility: Alterations have been made to the CMAT code such that it can now be run on the SGI Origin 2000 and Sun Microsystems SunFire V880 Servers at UCL's HiPerSPACE computing facility, and on the Hewlett Packard HP-UX machine at the Space Environment Centre in Colorado. This involved the inclusion of error trapping and significant improvements to the robustness, usability and portability of the code.

New Lower Boundary: In the new version of the CMAT model, the empirical MSISE90 model has been replaced by the NRLMSIS-00 empirical model, used to determine lower boundary seasonal geopotential height oscillations, lower boundary temperature and several start-up chemical constituent profiles.

Chapman Function: Numerical evaluation of the Chapman Grazing incidence function is now calculated using the expressions of *Smith and Smith* [1972]. These expressions are not only simpler and computationally more efficient than those previously used, but result in smoother profiles of major species photoionisation rates during the day-night transition.

Solar Spectrum and Thermospheric heating: In the first version of the CMAT model, thermospheric heating was calculated considering solar radiation in the 1.8 to 180nm range at 5nm spectral intervals. Heating due to solar UV, EUV and X-ray radiation is now calculated in the wavelength range 0.2 to 180nm at 1nm resolution or better. For wavelengths of 1.8nm and above, solar flux data is taken from the empirical SOLAR2000 model (*Tobiska* [2000]). For lower wavelengths, solar fluxes are taken

directly from the GLOW model (*Solomon and Abreu* [1989]). Variation of the solar flux input with solar activity was previously calculated by linearly interpolating between two spectra, one appropriate to low and one to high F10.7 cm flux values. Solar cycle variability is now accounted for by using interpolation between 12 different empirical solar spectra, appropriate to F10.7 values between 67 and 243.

Updated photoionisation and absorption cross sections: Major species photoionisation and absorption cross sections have been updated and tabulated in bins of 1nm or less. Combined with the finer resolution solar flux data, this has lead to better agreement between CMAT predicted temperatures and those given by NRLMSIS-00 under conditions of high and low solar activity during equinox and solstice. CMAT calculated F2 peak electron densities are now in better agreement with those given by IRI. Agreement between the magnitudes of the atomic oxygen 557nm green line volume emission rate calculated by CMAT and those given by the WINDII satellite has also been improved as a result of these updates.

Updated reaction rates and branching ratios: Recently published branching ratios and reaction rates have been collated, tested and included in the CMAT chemical scheme, along with several reactions that were previously omitted.

NO, NO_2 and $N(^4S)$ start up climatologies: One dimensional polynomials previously used to define start up profiles for thermospheric nitric oxide concentration have been replaced with observed number densities, derived from SNOE satellite data. The new start-up data sets include meridional and seasonal variability. $N(^4S)$ start up climatologies are now taken from NRLMSIS-00. This improves model run time by reducing the amount of time taken for a stable self-consistent solution to be reached.

Vibrational relaxation of NO(v=1) by O atoms: 5.3μm infrared radiative emission from excited nitric oxide is an important cooling mechanism in the thermosphere. The deactivation rate coefficient of vibrationally excited NO has been updated from 6.5 $x10^{-17}$ m^3s^{-1} to 2.4 $x10^{-17}$ m^3s^{-1} after *Dodd et al.* [1999].

Medvedev and Klaassen gravity wave scheme: The gravity wave parameterisation used in the original CMAT model is a hybrid *Lindzen-Matsuno* scheme, due to *Meyer* [1999a,b]. While able to reasonably recreate the effects of gravity wave breaking on the middle atmosphere wind structure, it is rather simplistic and ignores the effects of wave-wave interactions. The scheme is finely tuned meaning certain parameters must be adjusted in order gain a reasonable representation of the observed zonal wind structure at different times of the year. The parameterisation has also been shown to reduce the amplitude of the diurnal tide (*Harris et al.* [2002]). A new gravity wave scheme based on *Medvedev and Klaassen* [1995, 2000, 2003] has therefore been implemented. In this scheme, the damping of a given gravity wave component due to low frequency 'background' components is considered. An isotropic gravity wave spectrum is used, consisting of 60 waves with 15 frequencies each, in the 4 cardinal directions. Phase velocities range from approximately 2 to 60ms^{-1}. A self-consistent calculation of the vertical eddy diffusion profile has also been implemented, as described by *Medvedev and Klaassen* [2003].

7.1.2 MODELLING OF NITRIC OXIDE IN THE LOWER THERMOSPHERE

The development of a 1-dimentional photochemical model of nitric oxide was described in chapter 4. The model calculates NO in the lower thermosphere considering 37 ion-neutral and neutral-neutral chemical reactions. Under conditions of low and high solar and geomagnetic activity, reasonable agreement exists between NO profiles calculated by the model and those given by the empirical NOEM model (*Marsh et al.* [2004]).

A comparison of modelled NO density profiles using solar flux data derived from the SOLAR2000 empirical model and the *Hinteregger et al.* [1981] spectrum reveals a increase in equatorial NO densities of up to a factor of 4 when using the SOLAR2000 fluxes. The impact of changes in key branching ratios and reactions rates on calculated NO densities has been also tested and quantified.

By including the 1D photochemical model in the CMAT GCM, the large deficiency in calculated NO number densities described by *Harris* [2001] has been removed at all

latitudes. Reasonable agreement exists between CMAT calculated zonal mean NO number density profiles and those given by NOEM under conditions of low and high solar and geomagnetic activity. Discrepancies between high latitude NO densities predicted by NOEM and CMAT are thought to arise from the over simplistic representation of auroral energy inputs used in the CMAT model. The match between observed and modelled NO densities is considered sufficiently good that CMAT can be used to simulate the temporal and spatial variation in thermospheric NO densities under conditions of low and high solar and geomagnetic activity.

7.1.3 THE TEMPORAL AND SPATIAL RESPONSE OF NITRIC OXIDE AND THERMOSPHERIC TEMPERATURE TO HIGH GEOMAGNETIC ACTIVITY.

In chapter 5, three studies were performed. In the first, the CMAT model was used to determine the range of latitudes in which significant quantities of aurorally produced nitric oxide are present at the peak altitude of 110km. Under conditions of moderate solar and geomagnetic activity, aurorally produced NO is found to contribute to the NO concentration at latitudes poleward of 45° in the northern hemisphere and 30° in the south. Meridional winds act to transport NO from high to low latitudes, resulting in enhanced mid latitude NO densities. At low latitudes, NO densities are strongly influenced by the vertical diurnal tide.

In the second study, NO densities at 110km resulting from two 6-day CMAT simulations were compared. The first simulation included a 14-hour period of high geomagnetic activity, while the second used inputs appropriate to constant moderate auroral activity throughout the full 6-day period. Results revealed a complex spatial and temporal distribution of NO enhancements arising from the storm forcing. The magnitude of modelled NO enhancements is dependent on geographic location with respect to the auroral oval. Maximum increases in NO number density occur between 14 and 48 hours after relaxation of the auroral energy inputs, depending on location. For the conditions simulated in this study, high and mid latitude NO densities are enhanced for several days after relaxation of the storm.

Temperature enhancements resulting from the simulated storm show an equally complex distribution with location and time. Maximum temperature enhancements of about 250K (50K) and 370K (80K) are predicted for high (low) latitudes at altitudes of 130km and 300km respectively. The timescales for relaxation to pre-storm temperatures vary with altitude. At 120km, a reasonably steady decrease occurs over a period of about 60 hours. At altitudes of 180km and above, temperatures decrease rapidly when storm forcing is removed, reaching pre-storm levels around 36 hours later. Considerably greater relaxation times are predicted with the CTIP model. This is attributed to the omission of NO cooling which is calculated self consistently in the CMAT model.

In the final modelling study, the CMAT and CTIP models were used to simulate the geomagnetically active period from 23rd October to 3rd November 2003. Good agreement was found between FPI temperatures measured at Kiruna, northern Sweden, and those simulated by the models. CMAT data provided the best match with observations in both magnitude and variability. A shortfall in the model temperatures predicted for the very intense storm period from 29th to 31st October is attributed to shortcomings in the auroral energy inputs.

An assessment of the meridional distribution of CMAT temperature enhancements at 250km altitude revealed greater temperature increases in the southern hemisphere, as might be expected for the time of year. During the most intense period of geomagnetic activity, maximum temperatures of 1800K and 1600K were predicted for the southern and northern hemisphere high latitude regions respectively. Simulated equatorial temperatures reached a maximum of 1600K as a result of increases in the F10.7 solar activity proxy from 181 to a maximum of 275. The intense auroral and solar activity inputs to the model also results in large increases in the NO number density at 110km. The magnitude and location of NO density peaks are highly variable in space and time. Changes in solar flux inputs lead to large changes in low latitude NO densities. Location with respect to the auroral oval and transport of NO by meridional winds are also key factors in determining the NO distribution. CMAT predicts maximum NO concentrations of nearly 14 $x10^{14}m^{-3}$ between –10° and –30° latitude, 180° longitude in

the southern hemisphere. Peaks in NO density occur one to two days after the onset of high latitude storm forcing. Strong meridional winds act to transport aurorally produced NO all the way to the equator, resulting in maximum equatorial densities of 11.6 $x10^{14}m^{-3}$, occurring approximately 21 hours after the end of the most intense geomagnetic storm.

7.1.4 THE EFFECT OF GRAVITY WAVE DRAG PARAMETERISATIONS ON THE DIURNAL TIDE IN CMAT.

In chapter 6, the CMAT model was used to investigate the effects of parameterised gravity wave momentum deposition on the zonal wind, the diurnal tide and subsequently on the atomic oxygen green line volume emission rate. The impact of gravity wave drag on the MLT region, as calculated by the parameterisations of *Meyer* [1999a,b] and *Medvedev and Klaassen* [2000] were compared under conditions of low solar activity during northern spring equinox. Momentum sources from both gravity wave parameterisations acted to reduce the mid-latitude zonal jets, resulting in reasonable agreement between CMAT and HWM daily mean zonal winds. The magnitude and morphology of the low latitude daily mean zonal wind was in better agreement with UARS observed winds when the *Medvedev and Klaassen* [2000] scheme was used.

As suggested by *Harris et al.* [2002], gravity wave momentum deposition as parameterised by the *Meyer* [1999a,b] scheme resulted in a factor of 2-3 reduction in the amplitude of the diurnal tide in the MLT region. Use of the *Medvedev and Klaassen* [2000] parameterisation however resulted in a small increase in the diurnal tidal amplitude. Momentum deposition from both schemes acted to advance the phase of the diurnal tide such that the peak amplitude at a given height occurs at an earlier time of day.

The magnitude and morphology of the atomic oxygen 557.7nm volume emission rates modelled by CMAT using the *Medvedev and Klaassen* [2000] gravity wave scheme were in good agreement with observations from the WINDII instrument onboard the UARS satellite. The tidal forcing required to realistically simulate the diurnal structure

of the airglow is however too weak when the *Meyer* [1999a,b] scheme is used. Good agreement between the absolute magnitudes of modelled and observed green line emission rates are attributed to the updated solar spectrum and associated photochemical parameters now used in the CMAT GCM.

7.2 FUTURE MODEL IMPROVEMENTS

There are many potential areas for improvement and expansion of the CMAT model. Solar cycle variability could for example, be determined by the increasingly popular E10.7 solar activity proxy; thought to provide a more accurate representation of solar flux variability in the high-energy part of the spectrum (*Tobiska et al.* [2000], *Tobiska* [2001]). There are however, certain key physical processes that are at present, rather poorly represented in the model. Improved representation of these processes would increase the accuracy and expand the capabilities and application of the model enormously. It is these limitations that shall be discussed in the following sections.

7.2.1 HIGH LATITUDE ENERGY INPUTS

The current representation of auroral energy input in CMAT is based on two sets of statistical data, the electric field patterns derived by *Foster et al.* [1986] and the TIROS/NOAA particle precipitation patterns of *Fuller-Rowell and Evans* [1987]. These sets of data were derived from different sources and were taken at different times. As such, there is no correlation between the placement of the energetic particle distribution and the convection patterns, apart from that which arises statistically. By changing the position of high latitude inputs in the Coupled Thermosphere-Ionosphere Model (CTIM), *Schoendorf et al.* [1996] showed that dramatic changes in modelled high-latitude electron densities result from small changes in high latitude inputs. Geographic consistency between these two fundamental energy sources would be a significant improvement in the CMAT GCM.

Another problem with the high latitude forcing is that the statistically averaged nature of the auroral energy inputs leads to inaccuracies in the latitudinal extent of the auroral oval, particularly at times of high geomagnetic activity. *Fuller-Rowell et al.* [1994] pointed out that the observed auroral oval may be found up to 10 degrees equatorward

of the modelled oval. In the same study, it was noted that statistical blurring at high latitudes smoothes out the local response and reduces peak magnitudes in both electric field strength and particle precipitation. The coarse temporal and spatial scales of the CMAT grid restrict the level of detail to which highly variable auroral energy inputs can be represented. Increases in the CMAT grid resolution would, however, be of limited benefit without a decrease in the spatial scales of the high latitude inputs. Reducing the time step used for high latitude inputs in the CMAT model does improve authenticity of geomagnetic storm simulations, but results in large increases in model run time which can be prohibitive for long term studies.

Finally, the electric field data used in CMAT includes statistical convection patterns for very high and very low activity. The percentage occurrence of these high and low activity levels in the original data set is just a few percent for the low activity levels, and 1% for the highest activity level. This means the statistical convection patterns for the lowest and highest index levels are based on insufficient data be to statistically significant (*Foster et al.* [1986]).

It is clear that there are many shortfalls in the statistical models of high latitude energy inputs currently used in the CMAT model. Alternative sources of electric field and particle precipitation data are however limited. While it is possible to model specific periods using high latitude forcing derived from the assimilative mapping of ionospheric electrodynamics (AMIE) procedure (e.g. *Emery et al.* [1996]), this procedure is not available for extended periods of time and is therefore not suitable for general use in CMAT. Accurate specification of the driving forces associated with geomagnetic storms is clearly a requirement in the CMAT model but limitations in available data make perfect matching of storm energy and momentum inputs problematic.

7.2.2 PLANETARY WAVE FORCING

Planetary wave dissipation is currently simplistically represented in CMAT by Rayleigh drag. A previous lack of success in attempts to include lower boundary planetary wave forcing has been attributed to the coarse longitudinal resolution (*Harris*

[2001]). The NCAR TIME-GCM model has a longitudinal grid size of 5° and has been used to reasonably simulate planetary wave forcing by coupling it to a lower atmosphere model (*Liu and Roble* [2002]). A similar resolution in the CMAT model may enable realistic planetary wave dissipation to be simulated via lower boundary forcing with UARS geopotential height data. This would greatly improve representation of the winter upper stratosphere where planetary wave breaking has a strong impact on meridional circulation.

7.2.3 STRATOSPHERIC CHEMISTRY

Several catalytic cycles that are important in stratospheric ozone chemistry are currently not represented in CMAT. These cycles, involving Cl_x (Cl and CLO) or Br_x (Br and BrO) chemical species, are not included due to their high level of complexity. Any overestimate in stratospheric ozone resulting from this omission will affect the temperature of the stratosphere and thus the height of overlying pressure levels. One possible implication of this would be an overestimation of the altitudes at which maximum airglow is predicted to occur. *Harris* [2001] suggested that climatologies of Cl_x based on UARS satellite data could be included in CMAT. If possible, a similar solution could be implemented for Br_x, using observed or modelled distributions.

7.2.4 PHOTOELECTRON SPECTRA

As discussed in chapter 4, the current representation of secondary electron photoionisation in CMAT is based on the code of *Strickland and Meier* [1982], as described by *Fuller-Rowell* [1993]. Low latitude major species ionisation rates used in CMAT are notably lower than those calculated by the GLOW model, as presented by *Bailey et al.* [2002]. The shortfall is attributed to the limited photon energy spectra used in the *Strickland and Meier* model, which did not account for photons or primary photoelectrons with energies greater than 450eV. Below 100km, vertical profiles of the ratio of photoelectron ionisation to primary ionisation of the major species are either extrapolated or estimated. Profiles of secondary electron ionisation are currently invariant with solar activity and the process of Auger ionisation is not accounted for in CMAT.

Photoelectron ionisation plays an important role in determining ion and electron densities in the atmosphere and as such a more detailed and up to date calculation of the process would improve the CMAT model. Calculation of photoelectron spectra is computationally intensive and therefore not appropriate for inclusion in the CMAT model. A parameterisation suitable for implementation in CMAT could however be developed by utilizing models such as the GLOW model that calculate fluxes of energetic photoelectrons and auroral electrons including the effects of transport.

7.2.5 GRAVITY WAVE PARAMETERISATION

The two gravity wave parameterisations currently integrated into the CMAT model both use a relatively simple spectrum of gravity waves. Inclusion of a realistic seasonally varying anisotropic gravity wave spectrum would improve the representation of gravity wave momentum deposition in the CMAT model. This would enable processes affected by seasonal changes in MLT dynamics to be modelled with much greater accuracy. Orographic waves could also be included in the *Medvedev and Klaassen* [2000] gravity wave spectrum.

7.2.6 MODEL USABILITY AND EFFICIENCY

The size and complexity of the CMAT model make it extremely difficult for the inexperienced user to develop or run. Years of development in an uncontrolled environment have lead to inconsistencies in the code. For example numerous variable names are assigned multiple values in different routines. Key parameters are sometimes specified or calculated multiple times throughout the code making the tracing of errors time consuming and problematic. The horizontal resolution of the model is hard coded meaning a change to variable resolution would be non trivial. Inefficient structuring means the model run time is prohibitively long, restricting the study of the long time periods required for climate change studies. Perhaps the greatest problem with the existing version of CMAT is however, the lack of detailed documentation. All of these issues have been addressed in the CMAT2 model which is currently being developed.

7.2.7 LOW LATITUDE PLASMASPHERE MODEL

At latitudes between ±25°, CMAT O^+ density, and ion and electron temperatures are not self consistently calculated, but linearly interpolated from the boundary values of the high latitude code. Electron and molecular ion density are derived from these interpolated parameters. Inclusion of a self-consistent plasmasphere code such as that included in the CTIP model (*Millward et al.* [1996], *Fuller-Rowell et al.* [1996]) would improve the calculation of low latitude electron densities and may reduce the E region electron densities that are currently overestimated. Inclusion of the plasmasphere would also facilitate inter-hemispheric transport of charged particles.

7.3 FUTURE STUDIES

While there are almost endless possible studies that could be performed with the CMAT GCM, the details of the model make it more applicable to certain subject areas than others. The developments described in this thesis have extended the scope of the model such that previously unsuitable topics can now be explored. Some of the potential studies that have been made possible by the work in this thesis are outlined below.

7.3.1 DIURNAL AND SEASONAL VARIABILITY OF NO

In chapter 4, zonal mean profiles of nitric oxide number densities as calculated by the CMAT model were presented. This format was chosen as it allowed direct comparison with satellite and empirical model data, which is available only as a zonal mean for a single local time. It would be of interest to study the variability of NO on diurnal to seasonal timescales. A study by *Burns et al.* [2004] highlighted hemispheric differences in the thermal and compositional response of the atmosphere to geomagnetic forcing at solstice. This was attributed to differences in nitric oxide concentration in the summer and winter hemispheres. The CMAT model could be used to investigate these differences and their impact on the thermal and compositional structure of the geomagnetically disturbed atmospheric.

7.3.2 VIBRATIONAL RELAXATION OF NO(v = 1) AND $CO_2(01^10)$

Following excitation by collisions with O atoms, infrared emissions from excited CO_2 and NO result in cooling which plays an important role in determining the thermal budget of the thermosphere. There is still a great deal of uncertainty surrounding the vibrational deactivation rates associated with these emissions. Generally accepted experimental and observationally derived rates for the deactivation of CO_2 and NO vary from 1.0 to 9 x $10^{-12}cm^3s^{-1}$ and 1.9 x 10^{-11} to 6.5 x 10^{-11} cm^3s^{-1} respectively (*Sharma and Roble* [2002], *Kockarts* [1980]). Values currently used in the CMAT model are 3.5 x 10^{-12} cm^3s^{-1} for the deactivation of $CO_2(01^10)$, and 2.4 x10^{-11} cm^3s^{-1} for the deactivation of NO(v=1) (*Dodd et al.* [1999]). *Sharma and Roble* [2002] found that variation of these rates in a global mean model resulted in dramatic changes in the temperature and density structure of the thermosphere. Between them, the two rate coefficients have a major impact on the energetics and dynamics of the region. An investigation into the influence of these rate coefficients on the CMAT thermospheric structure could perhaps assist in refining the range of possible values and would be an interesting study.

7.3.3 DOWNWARD TRANSPORT OF NO_x INTO THE STRATOSPHERE

Numerous studies have investigated a possible coupling between the upper and lower atmosphere via the downward transport of odd nitrogen at high latitudes, from the thermosphere to mesospheric and even stratospheric altitudes (*Callis et al.*, [1996], *Callis and Lambeth* [1998], *Siskind et al.* [1997], *Randall et al.* [2001], *Callis et al.* [2002]). This is of particular interest after periods of enhanced geomagnetic activity when large concentrations of nitric oxide are produced. The result is enhanced concentrations of NO_x in the stratosphere that can contribute to the catalytic destruction of ozone. During polar night when aurorally produced NO is long-lived, NO_x can be transported over periods of weeks or months, meaning constituent effects may persist for some time after the initial production event (*Seppala et al.* [2004]). Despite numerous studies into the contribution of thermospheric NO to middle atmosphere NO_x concentrations (*Jackman et al.* [1995], *Callis et al.* [1996], *Siskind et*

al. [1997], *Vitt et al.* [2000]), no attempt has been made to model the temporal and spatial scales of NO descent in 3D GCMs.

In order to model this process correctly in CMAT, a reasonable representation of wintertime middle atmosphere circulation is required, including planetary wave induced meridional transport and associated downward species advection at high latitudes. Planetary wave mixing also affects the magnitude and duration of unmixed descent in winter which plays a key role in determining the amount of NO deposited in the springtime upper stratosphere (*Siskind et al.* [1997]). If planetary wave mixing is not included, enhanced NO remains confined to high latitudes. The inclusion of lower boundary planetary wave forcing in CMAT, as discussed above, would greatly improve the representation of odd nitrogen transport, and allow this topical coupling mechanism to be studied in detail.

7.3.4 FURTHER STUDIES OF THE OCTOBER 2003 GEOMAGNETIC STORM.

Modelling studies of periods of high solar and geomagnetic activity are particularly useful in broadening understanding of the energetic and dynamical processes associated with high energy inputs to the atmosphere. Compositional changes that result directly and indirectly from these dramatic energy inputs can often only be studied with models as the spatial and temporal coverage of observational data is generally very limited. The third study carried out in chapter 5 confirmed that the CMAT model is able to reasonably recreate thermospheric temperatures arising from periods of intense geomagnetic activity. The statistical representation of particle precipitation and convection electric fields does however restrict the ability of the model to realistically simulate the actual high latitude energy inputs that occurred over the period. The generic solar spectrum used restricts the authenticity of specific spectral features. This means that the intense UV and X-ray fluxes measured over the period (*Woods et al* [2004]) are not accurately represented. Updates to the thermospheric heating routine in CMAT mean that it is now possible to run the model with high resolution solar flux data, including temporal variability. As such it would be possible to simulate the October 2003 storm period using observational spectral intensity data, such as that measured by the TIMED, SOURCE and GOES satellites.

Similarly, empirical convection electric field patterns such as those measured by the Super Dual Auroral Radar Network (SuperDARN) could be implemented in the model, improving the spatial representation of high latitude energy inputs. Energetic particle precipitation fluxes from the NOAA-POES satellite could also be used to provide an accurate representation of high latitude particle and energy flux in the model. Another alternative would be to use the Assimilative Mapping of Ionospheric Electrodynamics (AMIE) procedure (*Emery et al.* [1996]) to specify high latitude ionospheric convection patterns and auroral particle precipitation. With these inputs, a detailed analysis of the energetic, dynamic and compositional changes resulting from this exceptionally active period may be assessed.

APPENDIX A: THE UPDATED CMAT CHEMICAL SCHEME

Bold type indicates a change or addition to the original chemical scheme. Reaction rates can be found in Table 3.2.

A.1 ODD OXYGEN - $O_x = O_3 + O$

$$\frac{dO_3}{dt} = \{(O)^2(O_2)k_{1a} + (O)(O_2)^2 k_{1b} + (O)(O_2)(N_2)k_{1c} + \mathbf{(O)(O_2)(M)k_2}\} -$$

$$O_3\{(J_{O3a} + J_{O3b}) + (H)a_2 + (OH)a_6 + (HO_2)a_{6b} + (O_3)(NO)b_4 + (NO_2)b_9 +$$

$$(O)k_3\}$$

$$\frac{dO}{dt} = \{(OH)^2 a_{16} + (H)(HO_2)a_{23c} + (N\,^4S)(O_2)b_7 + (N\,^2D)(O_2)b_{7b} + (N\,^2D)(NO)b_{6a} +$$

$$(N\,^4S)(NO)b_6 + J_{NO2}(NO_2) + J_{NO3a}(NO_3) + (J_{O3a} + J_{O3b})(O_3) + 2(J_{SRB,HZ} +$$

$$J_{SRC})(O_2) + \textit{ion\_prod\_O} - (O\,^1D)(H_2)a_{3b} - (O\,^1D)(H_2O)a_{1b} - (O\,^1D)(CH_4)c_{1a}\}$$

$$- O\{(H_2)a_3 + (OH)a_5 + (HO_2)a_7 + (H_2O_2)a_{31} + (CO)(M)a_{37} + (NO_2)b_3 +$$

$$(O)(CH_4)c_3 + (CLO)d_3 + (O_2^2)k_{1b} + (N_2)(O_2)k_{1c} + \mathbf{(O_2)(M)k_2} + (O_3)k_3 +$$

$$\textit{ion\_loss\_O}\} - (O)^2\{(O_2)k_{1a}\}$$

where

$$\textit{ion\_prod\_O} = (N^+)(O_2)\mathbf{R_4} + (O_2^+)(N\,^4S)R_8 + (O^+)(O_2)R_{10} + 2(O_2^+)(e^-)R_7 +$$

$$(NO^+)(e^-)R_2 + 0.33(\textit{aurqo2})$$

$$\textit{ion\_loss\_O} = (N^+)R_6 + (N_2^+)R_1 + Ji_O(1 + \textit{pe\_O(z)}) + \textit{aurqo}$$

where *aurqo2* and *aurqo* are the ionisation rates due to auroral sources.

pe_O(z) is the ratio of photoelectron ionisation to initial solar photon ionisation at altitude z as described in section 2.14.1.

$O(^1D)$ is assumed to be in photochemical equilibrium such that

$$\frac{d(O^1D)}{dt} = 0 = J_{O3a}(O_3) + J_{SRC}(O_2) + J_{H2Ob}(H_2O) - O(^1D)[(H_2O)a_{1b} + (H_2)a_{3b} + (CH_4)c_{1a} + (O_2)k_{4b} + (N_2)k_{4c}]$$

$$O(^1D) = \frac{J_{O3a}(O_3) + J_{SRC}(O_2) + J_{H2Ob}(H_2O)}{(H_2O)a_{1b} + (H_2)a_{3b} + (CH_4)c_{1a} + (O_2)k_{4b} + (N_2)k_{4c}}$$

A.2 ODD HYDROGEN - HO_X = H + OH + HO_2

The chemical production and loss of the odd hydrogen family members in CMAT are shown below

$$\frac{dHO_x}{dt} = 2[\ (H_2O)J_{H2Oa} + (O\ ^1D)(H_2)a_{3b} + (O^1D)(H_2O)a_{1b} + (H_2O_2)J_{H2O2}$$

$$+ (O\ ^1D)(CH_4)c_{1a} + (O)(H_2)a_3] - 2[\ (HO_2)(H)(a_{23b} + a_{23c}) + (OH)(HO_2)a_{17}$$

$$+ (HO_2)^2\ a_{27} + (OH)^2\ a_{16} + (OH)(CH_4)c_2]$$

Hydrogen peroxide is assumed to be in a state of photochemical equilibrium such that

$$\frac{dH_2O_2}{dt} = 0 = a_{27}(HO_2^{\ 2}) - (H_2O_2)[J_{H2O2} + a_{30}(OH) + a_{31}(O)]$$

A.3 ODD NITROGEN - NO_X = NO + NO_2

The chemical production and loss of the odd nitrogen family members in CMAT are shown below

$$\frac{dNO_x}{dt} = [\ (O_2)(N\ ^2D)b_{7b} + (O_2)(N\ ^4S)b_7 + (NO_3)(J_{NO3a} + J_{NO3b}) + (N\ ^4S)(OH)\mathbf{b_{10}}$$

$$+ \mathbf{2(N_2O)\,(O(^1D))b_{13}} + \mathbf{(N^+)(O)R_{4a}}\ ] - [\ (NO)J_{NO} + (NO)(N\,^4S)b_{6b}$$

$$+ (NO)(N\,_2D)b_{6a} + (NO)Ji_{NO} + (NO_2)(O_3)b_9]$$

$$\frac{dNO}{dt} = A_{NO} - B_{NO}(NO)$$

$$A_{NO} = (NO_2)(O)b_3 + (N\,^4S)(O_2)b_7 + (N\,^2D)(O_2)b_{7b} + (N\,^4S)(OH)b_{10} + (NO_2)J_{NO2}$$

$$+ \mathbf{2(N_2O)(O(^1D))b_{13}} + \mathbf{(NO_3)J_{NO3b}} + \mathbf{(N^+)(O)R_{4a}}$$

$$B_{NO} = (HO_2)a_{26} + (O_3)b_4 + (N\,^4S)b_6 + (N\,^2D)b_{6a} + (O_2^+)R_9 + J_{NO} + Ji_{NO} + \mathbf{(ClO)d_4}$$

$$\frac{dNO_2}{dt} = A_{NO2} - B_{NO2}(NO_2)$$

$$A_{NO2} = (HO_2)(NO)a_{26} + (NO)(O_3)b_4 + (NO_3)J_{NO3a} + \mathbf{(NO)(CLO)d_4}$$

$$B_{NO2} = (O)b_3 + (O_3)b_9 + J_{NO2} + (NO_3)b_{12}$$

Photochemical equilibrium is assumed for NO_3, therefore in the upper stratosphere

$$\frac{dNO_3}{dt} = 0 = (O_3)(NO_2)b_9 - NO_3[b_{12}(M)(NO_2) + J_{NO3}]$$

$$NO_3 = \frac{(O_3)(NO_2)b_9}{[b_{12}(M)(NO_2) + J_{NO3}]}$$

A.4 ATOMIC NITROGEN - N(^{4}S)

$$\frac{d(N^4S)}{dt} = ion\_pro_{N4S} + (NO)J_{NO} + (O)(N\,^2D)b_8 + (N\,^2D)b_{xx1} + (N\,^2D)(e^-)\mathbf{R_{13}}$$

$$-(N\,^4S)\{(O_2)b_7 + (NO)b_6 + (OH)b_{10} + \mathbf{(O_2^+)R_8} + \mathbf{J_{N4S}}\}$$

ion_pro_{N4S} is the ionospheric production of N^4S

$$ion\_pro_{N4S} = \ 0.25(Ji_{PN2}) + \mathbf{0.15}(NO^+)(e^-)R_2 + \mathbf{0.90(0.44}(N_2^+)(e^-)\mathbf{R_3}) + (O^+)(N_2)R_{11}$$

$$+ (N^+)(O_2)\mathbf{R_5} + (N^+)(O)R_6 + 0.40((1.34)Pe\_N_2(z)\ Ji_{N2})) + Ji_{N2}(N\_Np(z))$$

$$+ 0.24\ Pe\_N_2(z)\ ) + 0.5(aurqn2*0.24) + 0.25(aurqn2*1.34)$$

$Pe\_N_2(z)$ is the ratio of photoelectron ionisation to initial solar ionisation at altitude z, $N\_Np(z)$ the ratio of N^+ production to primary photon ionisation of N_2 (Ji_{N2}) at altitude z, *aurqn2* the ionisation rate due to auroral sources. Ji_{PN2} is the pre-dissociation rate of N_2 and refers to the process where incident photons excite N_2 molecules to *predissociation states* which decay by dissociation of the molecule rather than by radiation. The resulting atomic nitrogen is partitioned between ground state ($N(^4S)$) and excited state ($N(^2D)$) nitrogen. The partitioning factors used in this expression are due to *Fuller-Rowell* [1993].

A.5 ATOMIC NITROGEN - $N(^2D)$

$$\frac{d(N^2D)}{dt} = ion\_pro_{N2D} - (N^2D)[(O_2)b_{7b} + (NO)b_{6a} + (O)b_8 + b_{xx1} + (e^-)R_{13}]$$

ion_pro_{N2D} is the ionospheric production of N^2D

$$ion\_pro_{N2D} = (N_2^+)(O)R_1 + 0.75(Ji_{PN2}) + \mathbf{0.85}(NO^+)(e^-)\mathbf{R_2} + \mathbf{0.1}(\mathbf{2}(N_2^+)(e^-)\mathbf{R_3})$$

$$+ \mathbf{0.9}(\mathbf{0.56}(N_2^+)(e^-)\mathbf{R_3}) + 0.60((1.34)Pe\_N_2(z)\ Ji_{N2})) + 10Ji_{N2}(N\_Np(z)$$

$$+ 0.24\ Pe\_N_2(z)) + 0.5(aurqn2*0.24) + 0.75(aurqn2*1.34)$$

A.6 WATER VAPOUR – H_2O

$$\frac{d(H_2O)}{dt} = (OH)(OH)a_{16} + (OH)(HO_2)a_{17} + (OH)(H_2O_2)a_{30} + (OH)(H_2)a_{19}$$
$$+ (H)(HO_2)a_{23c} + (CH_4)(OH)c_2 - H_2O[J_{H2Oa} + J_{H2Ob}]$$

A.7 MOLECULAR HYDROGEN – H_2

$$\frac{d(H_2)}{dt} = (HO_2)(H)a_{23b} + (H_2O)J_{H2Ob} + (H)^2(M)a_{24} - H_2[(OH)a_{19} + (O^1D)a_{3b} + (O)a_3]$$

A.8 METHANE – CH_4

$$\frac{d(CH_4)}{dt} = -CH_4[(OH)c_2 + (O)c_3 + (O^1D)c_{1a}]$$

A.9 CARBON DIOXIDE – CO_2

$$\frac{d(CO_2)}{dt} = (CO)(OH)a_{36} + (CO)(O)(M)a_{37}$$

A.10 CARBON MONOXIDE – CO

$$\frac{d(CO)}{dt} = (CH_4)(O^1D)c_{1a} + (CH_4)(O)c_3 + (CH_4)(OH)c_2 - (CO)[(OH)a_{36} + (O)(M)a_{37}]$$

A.11 HELIUM – He

$$\frac{d(He)}{dt} = 0$$

REFERENCES

Akmaev R. A., V. A. Yudin, and D. A. Ortland (1997), SMLTM simulations of the diurnal tide: comparison with UARS observations, *Ann. Geophys.*, 15, 1187-1197.

Akmaev R. A. (2001), Simulation of large-scale dynamics in the mesosphere and lower thermosphere with the Doppler-spread parameterization of gravity waves: 2. Eddy mixing and the diurnal tide, *J. Geophys. Res.*, 106 (D1), 1205-1213.

Allen M., J. I. Lunine, and Y. L. Lang (1984), The vertical distribution of ozone in the mesosphere and lower thermosphere, *J. Geophys. Res.*, 89, 4841-4872.

Arnold, N. F., and T. R. Robinson (1998), Solar cycle changes to planetary wave propagation and their influence on the middle atmosphere circulation, *Ann. Geophys.*, 16, 69-76.

Arnold, N. F., and T. R. Robinson (2000), Amplification of the influence of solar flux variations on the winter stratosphere by planetary waves, *Space Science Reviews*, 94, Issue 1-2, 279–286.

Bailey, S. M., C. A. Barth, and S. C. Solomon (2002), A Model of Nitric Oxide in the lower Thermosphere, *J. Geophys. Res*, 107, 1206-1218.

Baker, D. N., C. A. Barth, K. E. Mankoff, S. G. Kanekal, S. M. Bailey, G. M. Mason, and J. E. Mazur (2001), Relationship between precipitating auroral zone electrons and lower thermopsheric nitric oxide densities: 1998-2000, *J. Geophys. Res.*, 106 (A11) 24,465-24,480.

Balachandran, N. K., and D. Rind (1995), Modelling the effects of UV variability and the QBO on the troposphere-stratosphere system. Part I: The middle atmosphere. *J. Clim.*, 8, 2058-2079.

Banks, P. M., and G. Kockarts (1973), Aeronomy. Part A, *Academic*, San Diago, Calif.

Bartels, J. (1949), The standardized Index Ks and the planetary index K_p, *IATME Bull* 126, 97, IUGG Publ., Office, Paris.

Barth, C. A., and A. F. Hildebrand (1961), The 5577Å airglow emission mechanism, *J. Geophys. Res.*, 66, 985.

Barth, C. A., D. W. Rusch, and A. I. Stewart (1973), The UV nitric-oxide experiment for Atmosphere Explorer, *Radio Sci.*, 8, 379–385.

Barth, C. A., W. K. Tobiska, D. E. Siskind, and D. D. Cleary (1988), Solar-Terrestrial coupling: Low latitude Thermospheric Nitric Oxide, *Geophys. Res. Lett.*, 15, 92-94.

Barth, C. A. (1990), Reference models for thermospheric NO, *Adv. Space. Res.*, V10, 6, 103-115.

Barth, C. A. (1992), Nitric Oxide in the Lower Thermosphere, *Planet. Space. Sci.*, 40, No 2/3, 315-336.

Barth, C. A. (1996), Reference models for thermospheric nitric oxide, 1994, *Adv. Space. Res.*, V18, 9-10, 179-208.

Barth, C. A., S. C. Solomon, S. M. Bailey, K. D. Mankoff, and A. W. Merkel (1999a), Global Images of Nitric Oxide in the Thermosphere, AGU Poster, SM41A-05.

Barth, C. A., S. M. Bailey, and S. C. Solomon (1999b), Solar-terrestrial coupling: Solar soft X-rays and thermospheric nitric oxide, *Geophys. Res. Lett.*, 26, 1251.

Barth, C. A., D. N. Baker, K. D. Mankoff, and S. M. Bailey (2001), The northern auroral region as observed in nitric oxide, *Geophys. Res. Lett.*, 28, 1463.

Barth, C. A., K. D. Mankoff, S. M. Bailey and S. C. Solomon (2003), Global Observations of Nitric Oxide in the Thermosphere, *J. Geophys. Res.*, 108 (A1), 1027.

Barth, C. A., and S. M. Bailey (2004), Comparison of a thermospheric photochemical model with Student Nitric Oxide Explorer (SNOE) observations of nitric oxide, *J. Geophys. Res.*, 109, A03304.

Bates, D. R. (1951), The temperature of the upper atmosphere, *Proc. Phys. Soc.*, London, 64B, 805.

Beagley, S. R., J. de Grandpre, J. M. Koshyk, N. A. McFarlane, and T. G. Shepherd (1997), Radiative dynamical climatology of the first generation Canadian middle atmosphere model, *Atmosphere-Ocean*, 35, 293-331.

Beig, G., et al. (2003), Review of mesospheric temperature trends, *Rev. Geophys.*, 41 (4), 1015.

Bilitza, D. (2001), International Reference Ionosphere 2000, *Radio Science* 36, 2, 261-275.

Brasseur, G., and S. Solomon (1986), Aeronomy of the middle atmosphere, *D. Reidel Publishing*, Second edition.

Brasseur, G. (1993), The response of the middle atmosphere to long-term and short-term solar variability: A two-dimensional model, *J. Geophys. Res.*, 98 (D12), 23,079–23,090.

Brasseur, G., M. H. Hitchman, S. Walters, M. Dymek, E. Falise, and M. Pirre (1990), An interactive chemical dynamical radiative two-dimensional Model of the middle atmosphere, *J. Geophys. Res.*, 95 (D5), 5639-5655.

Brewer, A. W. (1949), Evidence for a world circulation provided by measurements of helium and water vapour distribution in the stratosphere, *Q. J. Roy. Met. Soc.*, 75, 351.

Buonsanto Buonsanto, M. J., P. G. Richards, W. K. Tobiska, S. C. Solomon, Y.-K. Tung, and J. A. Fennelly (1995), Ionospheric electron densities calculated using

different EUV flux models and cross sections: Comparison with radar data, *J. Geophys. Res.*, 100(A8), 14,569–14,580.

Burns, A. G., T. L. Killeen, W. Wang, and R. G. Roble (2004), The solar-cycle-dependent response of the thermosphere to geomagnetic storms, *J. Atmos. Sol. Terr. Phys.*, 66, 1-14.

Burrage, M. D., W. R. Skinner, D. A. Gell, P. B. Hays, A. R. Marshall, D. A. Ortland, A. H. Manson, S. J. Franke, D. C. Fritts, P. Hoffmann, C. McLandress, R. Niciejewsky, F. J. Schmidlin, G. G. Shepherd, D. Singer, T. Tsuda, and R. A. Vincent (1996), Validation of mesosphere and lower thermosphere winds from the high resolution Doppler imager on UARS, *J. Geophys. Res.*, 101 (D6), 10365-10392.

Callis, L. B., D. N. Baker, M. Natarajan, J. B. Blake, R. A. Mewaldt, R. S. Selesnick, and J. R. Cummings (1996), A 2-D model simulation of downward transport of NO_y into the stratosphere: Effects on the 1994 austral spring O_3 and NO_y, *Geophys. Res. Lett.*, 23 (15), 1905–1908.

Callis, L. B., and J. D. Lambeth (1998), NO_y formed by precipitating electron events in 1991 and 1992: Decent into the stratosphere as observed by ISAMS, *Geophys. Res. Lett.*, 25(11), 1875–1878.

Callis, L. B., M. Natarajan, J. D. Lambeth, and D. N. Baker (1998), Solar atmospheric coupling by electrons (SOLACE), 2, Calculated stratospheric effects of precipitating electrons, 1979-1988, *J. Geophys. Res.*, 103 (D21), 28, 421–28,438.

Callis, L. B., M. Natarajan, and J. D. Lambeth (2001), Solar-atmospheric coupling by electrons (SOLACE) 3. Comparison of simulations and observations, 1979-1997, issues and implications, *J. Geophys. Res.* 106 (D7), 7523-7539.

Callis, L. B., M. Natarajan, and J. D. Lambeth (2002), Observed and calculated mesospheric NO, 1992–1997, *Geophys. Res. Lett.*, 29 (2), 1030.

Chapman, S., and T. G. Cowling (1952), The mathematical theory of non-uniform gases, Chapters 5, 8, 14, *Cambridge Univ. Press*, New York and London.

Chapman, S., and R. S. Lindzen (1970), Atmospheric tides, *D. Reidel Publishing Co.*, Dodrecht.

Charney, J. G., and P. G. Drazin (1961), Propogation of planetary-scale disturbances from the lower into the upper atmosphere, *J. Geophys. Res.*, 66, 83-109.

Charney, J. G., R. Fjörtoft, and J. von Neumann (1950), Numerical integration of the barotropic vorticity equation, *Tellus*, 2, 237-54.

Cleary, D. D. (1986), Daytime high-latitude rocket observations of the NO γ, δ, and ε bands, *J. Geophys. Res.*, 91, 11,337.

Codrescu, M. V., T. J. Fuller-Rowell, R. G. Roble, and D. S. Evans (1997), Medium energy particle precipitation influences on the mesosphere and lower thermosphere, *J. Geophys. Res,* 102 (A9), 19,997-19,987.

Codrescu, M. V., T. J. Fuller-Rowell, J. C. Foster, J. M. Holt, and S. J. Cariglia (2000), Electric field variability associated with the Millstone Hill electric field model, *J. Geophys. Res.*, 105 (A3), 5265-5273.

Colgrove, F. D., Johnson, F. S., and Hanson W. B. (1966), Atmospheric composition in the lower thermosphere, *J. Geophys. Res.*, 84, 811-822.

Crutzen, P. J. (1975), A two-dimensional photochemical model of the atmosphere below 55km, Proceeding of the fourth international conference on climatic impact assessment program, U.S. department of transportation, Washington D.C.

Dalgarno, A., and F. J. Smith (1962), The thermal conductivity and viscosity of atomic oxygen, *Planet. Space. Sci.*, 9, 1-2.

DeMore, W. B., S. P. Sander, C. J. Howard, A. R. Ravishankara, D. M. Golden, C. E. Kolb, R. F. Hampson, M. J. Kurylo, and M. J. Molina (1997), Chemical kinetics and

photochemical data for use in stratospheric modeling, Evaluation No. 12, *NASA Jet Propulsion Laboratory*, JPL Publication 97-4.

Dickinson, R. E., E. C. Ridley, and R. G. Roble (1981), A three-dimensional general circulation model of the thermosphere, *J. Geophys. Res.*, 86, 1499-1512.

Dobson, G. M. G. (1956), Origin and distribution of polyatomic molecules in the atmosphere, *Proc. Roy. Soc. Lond.* A., 236, 187.

Dodd, J. A., R. B. Lockwood, E. S. Hwang, S. M. Miller, and S. J. Lipson (1999), Vibrational relaxation of NO(v=1) by oxygen atoms, *J. Chem. Phys.* 111, 3498-3507.

Donnelly, R. F. and Pope, J. H. (1973), The 1-3000Å solar flux for a moderate level of solar activity for use in modeling the ionosphere and upper atmosphere, *NOAA Technical Report*, ERL 276, Environ. Res. Lab., Boulder.

Duff, J. W. and, R. D. Sharma (1997), Quasiclassical trajectory study of NO vibrational relaxation by collisions with atomic oxygen, *J. Chem. Soc., Faraday Trans.* Atmospheric chemistry special issue, 93, 2645-2649.

Duff, J. W., H. Dothe, and R. D. Sharma (2003), On the Rate Coefficient of the $N(^2D)+O_2 \rightarrow NO+O$ Reaction in the Terrestrial Thermosphere, *Geophys. Res. Lett.*,30, 5, 1259-1263.

Emery, B. A., et al. (1996), Assimilative mapping of ionospheric electro-dynamics in the thermosphere-ionosphere general circulation model comparisons with global ionospheric and thermospheric observations during the GEM/SUNDIAL period of March 28 - 29, 1992, *J. Geophys. Res.*, 101, 26,681-26,696.

Fennelly, J. A., and D. G. Torr (1992), Photoionization and Photoabsorption Cross Sections of O, N2 O2, and N for Aeronomic Calculations, *At. Data and Nucl. Data Tables*, 51, 321.

Fernando, R. P., and I. W. M. Smith (1979), Vibrational relaxation of NO by atomic oxygen, *Chem. Phys. Lett.*, 66, 218-222.

Field, P. R., H., Rishbeth, R. J., Moffett, D. W., Idenden, T. J. Fuller-Rowell, G. H. Millward, and A. D. Aylward (1998), Modelling composition changes in F-layer storms, *J. Atmos. Sol. Terr. Phys.*, 60, 5, 523-543.

Forbes, J. M. (1982a), Atmospheric tides. I - Model description and results for the solar diurnal component. II - The solar and lunar semidiurnal components, *J. Geophys. Res.*, 87, 5222-5252.

Forbes, J. M. (1982b), Atmospheric tides 2. The solar and lunar semidiurnal components, *J. Geophys. Res.*, 87 (A7), 5241-5252.

Forbes, J. M, J. Gu, and S. Miyahara (1991), On the interactions between gravity waves and the diurnal propagating tide, *Planet. Space Sci.*, 39 (9), 1249-1257.

Fomichev, V. I., and G. M. Shved (1985), Paramaterization of the radiative flux divergence in the 9.6 micro-m O3 band, *J. Atmos. Terr. Phys.*, 47, 11, 1037-1049.

Fomichev, V. I., J. P. Blanchet, and D. S. Turner (1998), Matrix parameterisation of the 15μm band cooling in the middle and upper atmosphere for variable CO_2 concentration, *J. Geophys. Res.*, 103 (D10), 11,505-11528.

Foster, J. C., J. M. Holt, R. G. Musgrove, and D. S. Evans (1986), Ionospheric convection associated with discrete levels of particle precipitation, *Geophys. Res. Lett.*, 13, 656.

Fritts, D. C., and M. J. Alexander (2003), Gravity wave dynamics and effects in the middle atmosphere, *Rev. Geophys.*, 41(1), 1003.

Fritts, D. C., and J. R. Isler (1994), Mean motion and tidal and two-day structure and variability in the mesosphere and lower thermosphere over Hawaii, *J. Atmos. Sci*, 51, 2145-2164.

Fritts, D. C., and W. Lu (1993), Spectral estimates of gravity wave energy and momentum fluxes. Part II: parameterization of wave forcing and variability. *J. Atmos. Sci,*. 50, 3695-3713.

Fritts, D. C., and T. E. VanZandt (1993), Spectral estimates of gravity wave energy and momentum fluxes. Part I: energy dissipation, acceleration, and constraints. *J. Atmos. Sci,*. 50, 3685-3694.

Fritts, D. C., and R. A. Vincent (1987), Mesospheric momentum flux studies at Adelaide, Australia: Observations and a gravity wave-tidal interactions model, *J. Atmos. Sci.*, 44, 605-619.

Fujiwara, H., S. Maeda, H. Fukunishi, T. J. Fuller-Rowell, and D. S. Evans (1996), Global variations of thermospheric winds and temperatures caused by substorm energy injection, *J. Geophys. Res.*, 101(A1), 225–240.

Fuller, E. N., P. D. Schettler, and J. C. Giddings (1966), *Industrial and engineering chemistry*, 58, 19.

Fuller-Rowell, T. J., and D. Rees (1980), A three-dimensional time-dependant global model of the thermosphere, *J. Atmos. Sci*, 37, 2545-2567.

Fuller-Rowell, T. J. (1981), A three-dimensional time-dependant global model of the thermosphere, *Ph.D. Thesis, University of London.*

Fuller-Rowell, T. J., and D. Rees (1983), Derivation of a conservation equation for mean molecular weight for a two-constituent gas within a three-dimensional time-dependent model of the thermosphere, *Planet. Space. Sci.*, 31, 1209.

Fuller-Rowell, T. J. (1984), A two dimensional, high-resolution, nested-grid model of the thermosphere: 1 Neutral response to an electric field "Spike", *J. Geophys. Res.*, 89 (A5), 2971-2990.

Fuller-Rowell, T. J., and D. S. Evans (1987), Height-integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data, *J. Geophys. Res.*, 92, 7606.

Fuller-Rowell, T. J. (1993), Modelling the solar cycle change in nitric oxide in the thermosphere and upper mesosphere, *J. Geophys. Res.*, 98, 1571-1580.

Fuller-Rowell, T. J., M. V. Codrescu, R. J. Moffett, and S. Quegan (1994), Response of the thermosphere and ionosphere to geomagnetic storms. *J. Geophys. Res.*, 99 (A3), 3893-3914.

Fuller-Rowell, T. J., D. Rees, S. Quegan, R. J. Moffett, M. V. Codrescu, and G. H. Millward (1996a), A coupled thermosphere ionosphere model (CTIM), Solar terrestrial energy program (STEP), handbook of ionospheric models, edited by R.W. Schunk.

Fuller-Rowell, T. J., M. V. Codrescu, H. Rishbeth, R. J. Moffett, and S. Quegan (1996b), On the seasonal response of the thermosphere and ionosphere to geomagnetic storms, *J. Geophys. Res.*, 101 (A2), 2343–2354.

Fuller-Rowell, T. J., and D. Rees (1996), Numerical Simulations of the Distribution of Atomic Oxygen and Nitric Oxide in the Thermosphere and Upper Mesosphere. In CIRA 1986, Vol. III, "Trace Constituents of the Middle and Upper Atmosphere", *Adv Sp. Res.* 18, 9-10, 255-305.

Fuller Rowell, T. J., M. C. Codrescu, and P. Wilkinson (2000), Quantitative modeling of the ionospheric response to geomagnetic activity, *Ann. Geophys.*, 18, 766-781.

Galand, M., R. G. Roble, and D. Lummerzheim (1999), Ionization by energetic protons in Thermosphere-Ionosphere Electrodynamics General Circulation Model, *J. Geophys. Res.*, 104 (A12), 27,973-27,989.

Garcia, R. R., and S. Solomon (1983), A numerical model of the zonally averaged dynamical and chemical structure of the middle atmosphere, *J. Geophys. Res.*, 88 (C2), 1379-1400.

Geller, M. A., (1983), Dynamics of the middle atmosphere, *Space Science Rev.*, 34, 359.

Gerard, J. C., and C. A. Barth (1977), High-latitude nitric oxide in the lower thermosphere, *J. Geophys. Res.*, 82, 674-680.

Greer, R. G. H., D. P. Murtagh, I. C. McDade, P. H. G. Dickenson, L. Thomas, D. B. Jenkins, J. Stegman, E. J. Llewellyn, G. Witt, D. J. Mackinnon, and E. R. Williams, (1986), ETON 1 : a database pertinent to the study of energy transfer in the oxygen nightglow. *Planet. Space Sci.*, 34, 771.

Hagan, M. E., J. M. Forbes, and F. Vial (1995), On modelling migrating solar tides, *Geophys. Res. Lett.*, 22, 893-896.

Hagan M. E. (1996), Comparative effects of migrating solar sources on tidal signatures in the middle and upper atmosphere, *J. Geophys. Res.*, 101, 21,213-21,222.

Hagan, M. E. (1998), The Global Scale Wave Model homepage, *http://www.hao.ucar.edu/public/research/tiso/gswm/gswm.html*

Haigh, J. D. (1996), The impact of solar variability on climate, *Science*, 272, 981-984.

Haigh, J. D. (1999), Modelling the impact of solar variability on climate, *J. Atmos. Terr. Phys.*, 61, 63-72.

Hamilton, K. P., (1997), The role of parameterized drag in a troposphere-stratosphere-mesosphere general circulation model. In: Hamilton, K. (Ed.), *Gravity Wave Processes: Their Parameterization in Global Climate Models. Springer*, Berlin, pp. 337-350.

Harris, M. J. (2001), A new coupled middle atmosphere and thermosphere circulation model: Studies of dynamic, energetic and photochemical coupling in the middle and upper atmosphere, PhD. Thesis, *University of London.*

Harris, M. J., N. F. Arnold, and A. D. Aylward (2002), A Study into the effect of the diurnal tide on the structure of the background mesosphere and thermosphere using the new coupled middle atmosphere and thermosphere (CMAT) general circulation model, *Ann. Geophys.*, 20, 225-235.

Harwood, R. S., and J. A. Pyle (1975), A two-dimensional mean circulation model for the atmosphere below 80km, *Quart. Jour. Met. Soc.*, 101, 723-747.

Haurwitz, B. (1961), Frictional effects and the meridional circulation of the mesosphere, *J. Geophys. Res.*, 66, 2381.

Hays, P. B., D. L. Wu, and the HRDI Science Team (1994), Observations of the diurnal tide from *Space, J. Atmos. Sci.*, 51, 3077-3093.

Hedin, A. E. (1991), Extension of the MSIS Thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, 96 (A2), 1159-1172.

Hedin, A. E., E. L. Flemming, A. H. Manson, F. J. Schmidlin, S. K. Avery, and S. J. Franke (1993), Empirical wind model for the middle and lower atmosphere 1: Local time average, *NASA Tech. Memo. NASA TM-104581,* 85.

Henke, B. L., E. M. Gullikson, and J. C. Davis (1993), *Atomic Data and Nuclear Data Tables*, V. 54 No. 2.

Hickey Hickey, M. P., and R. L. Walterscheid (1994), Wave-modified mean exothermic heating in the mesopause region, *Geophys. Res. Lett.*, 21 (22), 2413–2416.

Hines, C. O. (1960), Internal gravity waves at ionospheric heights, *Canad. J. Phys.*, 38, 1441-1481.

Hines, C. O. (1997a), Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere Part 1: Basic formulation. *J. Atmos. Sol. Terr. Phys.*, 59, 371-386.

Hines, C. O. (1997b), Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere Part 2: broad and quasi monochromatic spectra and implementation. *J. Atmos. Sol. Terr. Phys.*, 59, 387-400.

Hinteregger, H. E. (1970), The extreme ultraviolet solar spectrum and its variation during a solar cycle, *Ann. Geophys*, 26, 547.

Hinteregger, H. E., K. Fukui, and B. R. Gilson (1981), Observational, reference and model data on solar EUV, from measurements on AE-E, *Geophys. Res. Lett.*, 8, 1147.

Holton, J. R. (1975), The dynamic meteorology of the stratosphere and mesosphere, Meteor. Monog. 15 (37), *Amer. Met. Soc.*, MA.

Holton, J. R. (1982), The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere, *J. Atmos. Sci.*, 39, 791-799.

Houghton, J. T. (1978), The stratosphere and mesosphere, *Q. J. Royal. Met. Soc.* 104, 1-29.

Houghton, J. T. (2002), The physics of atmospheres, 3rd edition, Cambridge U.K.; New York, *Cambridge Univ. Press.*

Jackman, C. H., M. C. Cerniglia, J. E. Nielsen, D. J. Allen, J. M. Zawodny, R. D. McPeters, A. R. Douglass, J. E. Rosenfield, and R. B. Rood (1995), Two-dimensional and three-dimensional model simulations, measurements, and interpretation of the influence of the October 1989 solar proton events on the middle atmosphere, *J. Geophys. Res.*, 100 (D6), 11,641–11,660.

Jones, R. A., and M. H. Rees (1973), Time dependent studies of the aurora, 1, Ion density and composition, *Planet. Space Sci.*, 21, 559.

Khol, H., and J. W. King (1967), Atmospheric winds between 100 and 700km and their effects on the ionosphere, *J. Atm. Sol. Terr. Phys.*, 29, 1045-1062.

King, W. T. and B. Crawford Jr. (1972), The integrated intensity of the Nitric Oxide fundamental band, *J. Quant. Spectrosc. Radiat. Transfer*, 12, 443-447.

Kley, D., G. M., Lawrence, and E. J., Stone (1977), The yield of $N(^2D)$ atoms in the dissociative recombination of NO^+, *J. Chem. Phys*, 66, 9, 4157-4165.

Kockarts, G. (1980), Nitric Oxide Cooling in the terrestrial Thermosphere, *Geophys. Res. Lett.*, 7, No. 2, 137-140.

Krishna Kumar, C., P. K. Swaminathan, D. E. Anderson, J.-H. Yee, M. R. Gunson, and M. C. Abrams (1995), ATMOS/ATLAS 1 measurements of thermospheric and mesospheric nitric oxide, *J. Geophys. Res.*, 100 (D8), 16,839–16,846.

Langel, R. A. (1992), International Geomagnetic Reference Field, 1991 Revision, *Eos Trans.*, AGU, 73, 182.

Lean, J. (1990), A Comparison of Models of the Sun's Extreme Ultraviolet Irradiance Variations, *J. Geophys. Res.*, 95 (A8), 11,933-11,944.

Levin, E., H. Partridge, and J. R. Stallcop (1990), Collision integrals and high temperature transport properties for N-N O-O and N-O, *Journal of Thermophysics and Heat Transfer*, 4, 469-477.

Lindzen, R. S. (1981), Turbulence and stress due to gravity wave and tidal breakdown, *J. Geophys. Res.*, 86 (C10), 9707-9714.

Liu, H., and M. E. Hagan (1998), Local heating/cooling of the mesosphere due to gravity wave and tidal coupling, *Geophys. Res. Lett.*, 25 (15), 2941–2944.

Liu, H., M. E. Hagan, and R. G. Roble (2000), Local mean state changes due to gravity wave breaking modulated by the diurnal tide, *J. Geophys. Res.*, 105 (D10), 12,381–12,396.

Liu, H., and R. G. Roble (2002), A study of a self-generated stratospheric sudden warming and its mesospheric-lower thermospheric impacts using the coupled TIME-GCM/CCM3, *J. Geophys. Res.*, 107 (D23), 4695.

London, J. (1980), In Proceedings of the Nato Advanced Institute on Atmospheric Ozone (Portugal), *U.S. Dept. of Transportation, FAA – Washington, D. C., USA* – No.FAA-EE-80-20.

Lopez-Puertas, M., and F. W. Taylor (2001), Non-Local Thermodynamic Equilibrium, *World Scientific Publishing Company, Incorporated.*

Maeda, S., T. J. Fuller-Rowell, and D. S. Evans (1989), Zonally averaged dynamical and compositional response of the thermosphere to auroral activity during September 18-24, 1984, *J. Geophys. Res.*, 94, 16 869-16,883.

Maeda, S., T. J. Fuller-Rowell, and D. S. Evans (1992), Heat budget of the thermosphere and temperature variations during the recovery phase of a geomagnetic storm, *J. Geophys. Res.*, 97 (A10), 14,947-14,957.

Magnotta, F., and H. S. Johnston (1980), Photodissociation quantum yields for the NO_3 free radical, *Geophys. Res. Lett.*, 7, 769.

Manson, A. H., C. E. Meek, J. Koshyk, S Franke, D. C. Fritts, D. Riggin, C. M. Hall, W. K. Hocking, J. MacDougall, K. Igarashi, R. A. Vincent (2002), Gravity wave activity and dynamical effects in the middle atmosphere (60-90km): observations from an MF/MLT radar network, and results from the Canadian Middle Atmosphere Model (CMAM), *J. Atmos. Sol. Terr. Phys.*, 64, 65-90.

Marsh, D. A., and J. M. Russell III (2000), A tidal explanation for the sunrise/sunset anomaly in HALOE low-latitude nitric oxide observations, *Geophys. Res. Lett.*, 27, 19, 3197-3200.

Marsh, D. A., and R. Roble (2002), TIME-GCM simulations of lower-thermospheric nitric oxide seen by the halogen occultation experiment, *J. Atmos. Sol. Terr. Phys.*, 64, 889-895.

Marsh, D. R., S. C. Solomon, and, A. E. Reynolds (2004), Empirical model of nitric oxide in the lower thermosphere, *J. Geophys. Res.*, 109, A07301.

Matsuno, T. (1982), A quasi one-dimensional model of the middle atmosphere circulation interacting with internal gravity waves, *J.Meteor.Soc.Japan*, 60, 215-226.

Mayr, H. G., J. G. Mengel, K. L. Chan, and H. S. Porter (2001), Mesosphere dynamics with gravity wave forcings: Part I. Diurnal and semi-diurnal tides, *J. Atmos. Sol. Terr. Phys.*, 63, 1851-1864.

McDade I. C, D. P. Murtagh, R. G. H. Greer, P. H. G. Dickenson, G. Witt, J. Stegman, E. J. Llewellyn, L. Thomas, and D. B. Jenkins (1986), ETON 2: quenching parameters for the proposed precursors of $O_2(b^1\Sigma_g^+)$ and $O(^1S)$ in the terrestrial nightglow, *Planet. Space Sci.*, 34, 789.

McLandress, C., Y. Rochon, G. G. Shepherd, B. H. Solheim, G. Thuillier, and F. Vial (1994), The meridional wind component of the thermospheric tide observed by WINDII on UARS, *Geophys. Res. Lett.*, 21, 2417-2420.

McLandress, C. G. G. Shepherd, B. H. Solheim, M. D. Burrage, P. B. Hays, and W. R. Skinner (1996a), Combined mesosphere/thermosphere winds using the WINDII and HRDI data from the Upper Atmosphere Research Satellite, *J. Geophys. Res.*, 101 (D6), 10,441-10,453.

McLandress, C. G. G. Shepherd, B. H. Solheim (1996b), Satellite observations of thermospheric tides: Results from the Wind Imaging Interferometer on UARS, *J. Geophys. Res.*, 101 (D2), 4093-4114.

McLandress, C. (1998), On the importance of gravity waves in the middle atmosphere and their parameterizations in general circulation models, *J. Atmos. Sol. Terr. Phys.*, 60, 1357-1383.

Medvedev, A. S. and G. P Klaassen (1995), Vertical evolution of gravity wave spectra and the parameterization of associated wave drag. *J. Geophys. Res,* 100, 25,841-25,853.

Medvedev, A. S., G. P Klaassen, and B. A. Bolville (1997), The parameterization of gravity wave drag based on the nonlinear diffusion of wave spectra. In: Hamilton, K (Ed.), *Gravity Wave Processes: Their Parameterization in Global Climate models*, Springer, Berlin, pp. 309-325.

Medvedev, A. S., and G. P Klaassen (2000), Parameterization of gravity wave momentum deposition based on nonlinear wave interactions: basic formulation and sensitivity tests, *J. Atmos. Sol. Terr. Phys.*, 62, 1015-1033.

Medvedev, A. S., and G. P Klaassen (2003), Thermal effects of saturating gravity waves in the atmosphere, *J. Geophys. Res,* 108 (D2), 4040.

Merkel, A. W., C. A. Barth, and S. M. Bailey (2001), Altitude determination of ultraviolet measurements made by the Student Nitric Oxide Explorer, *J. Geophys. Res.*, 106 (A12), 30, 283–30,290.

Meyer, C. (1999a), Gravity wave interactions with the diurnal propagating tide, *J. Geophys. Res.*, 104 (D4), 4223-4239

Meyer, C. (1999b), Gravity wave interactions with mesospheric planetary waves: A mechanism for penetration into the thermosphere-ionosphere system, *J. Geophys. Res.*, 104, 28,181-28,196.

Millward, G., R.J. Moffett, S. Quegan, and T.J. Fuller-Rowell (1996) A Coupled Thermosphere-Ionosphere-Plasmasphere Model (CTIP), *Solar Terrestrial Energy Program (STEP) Handbook*, editor R. W. Schunk.

Miyahara, S., and J.M. Forbes (1991), Interactions between gravity waves and the diurnal tide waves in the lower thermosphere, *J. Meteorol. Soc. Jpn.*, 69, 523-531.

Mlynczak, M. G., and S. Solomon (1991a), Middle atmosphere heating by exothermic chemical reactions involving odd hydrogen species, *Geophys. Res. Lett.*, 18, 37-40.

Mlynczak, M. G., and S. Solomon (1991b), Reply, *Geophys. Res. Lett*, 18, 1793-1794.

Mlynczak, M. G., and S. Solomon (1993), A detailed study of heating efficiency in the middle atmosphere, *J. Geophys. Res.*, 98, 10,517-10,541.

Mlynczak, M., et al. (2003), The natural thermostat of nitric oxide emission at 5.3 μm in the thermosphere observed during the solar storms of April 2002, *Geophys. Res. Lett.*, 30 (21), 2100.

Müller-Wodarg, I. (1997), Modelling perturbations propagating through the mesopause into the Earth's upper atmosphere, Ph.D. Thesis, *University of London.*

Müller-Wodarg, I., A. D. Aylward, and T. J. Fuller-Rowell (2001), Tidal oscillations in the thermosphere: a theoretical investigation of their sources, *J. Atm. Sol. Terr. Phys.*, 63, 899-914.

Müller-Wodarg, I.C.F, R.V. Yelle, M. Mendillo, L.A. Young, and A.D. Aylward (2000), The thermosphere of Titan simulated by a global three-dimensional time-dependent model, *J. Geophys. Res.*, 105 (A9), 20,833-20,856.

Murgatroyd, R. J., and F. Singleton, Possible meridional circulation in the stratosphere and mesosphere (1961), *Q. J. Jour. Met. Soc.* 87, 125.

Murtagh, D. P., G. Witt, J. Stegman, I. C. McDade, E. J. Llewellyn, F. Harris, and R. G. H Greer (1990), An assessment of proposed $O(^1S)$ and $O_2(b^1\Sigma_g^+)$ nightglow excitation parameters, *Planet. Space Sci.*, 38, 43-53.

NASA (2000), Studying Earth's Environment From Space, http://www.ccpo.odu.edu/SEES/index.html

Newell, P. T., K. M. Lyons, and C.-I. Meng (1996), A large survey of electron acceleration events, *J. Geophys. Res.*, 101, 2599.

Nicolet , M., and Aiken, A. C. (1960), The formation of the region D of the ionosphere, *J. Geophys. Res.*, 65, 1469-1483.

Nicolet, M. (1984), Photodissociation of molecular oxygen in the terrestrial atmosphere : Simplified numerical relations for the spectral range of the Schumann-Runge Bands, *J. Geophys. Res.*, 89, 2573-2582.

Nicolet, M. (1979), Photodissociation of nitric oxide in the mesosphere and stratosphere: Simplified numerical relations for atmospheric model calculations, *Geophys. Res. Lett.*, 6, 866

Oshio, M., R. Madea, and H. Sakagami (1966), Height distribution of local photoionization efficiency, *J. Radio. Res. Lab. Jpn.*, 13, 245-577.

Pancheva D., N. J. Mitchell, M.E. Hagan, A.H. Manson, C.E. Meek, Y. Luo, C. Jacobi, D. Kurschner, R. R. Clark, W. K. Hocking, J. MacDougall, G. O. L. Jones, R. A. Vincent, I. M. Reid, W. Singer, K. Igarashi, G. I. Fraser, T. Nakamura, T. Tsuda, Y. Portnyagin, E. Merzlyakov, A. N. Fahrutdinova, A. M. Stepanov, L. M. G. Poole, S. B. Malinga, B. L. Kashcheyev, A. N. Oleynikov, and D. M. Riggin (2002), Global-scale tidal structure in the mesosphere and lower thermosphere during PSMOS campaign of June-August 1999 and comparisons with the Global Scale Wave model, *J. Atm. Sol. Terr. Phys.*, 64, 8-11, 1011-1035.

Picone, J.M., A.E. Hedin, D.P. Drob, and A.C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys. Res.*, 107 (A12), 1468.

Portnyagin Y. I., T. V. Solovjova, and D. Y Wang (1999), Some results of comparison between the lower thermosphere zonal winds as seen by the ground-based radars and WINDII on UARS, *Earth. Plan. Sp.*, 51, 701-709.

Quegan, S., G. J. Bailey, R. J. Moffett, R. A. Heelis, T. J. Fuller-Rowell, D. Rees, and A.W. Spiro (1982), A theoretical study of the distribution of ionization in the high-latitude ionosphere and the plasmasphere: First results of the mid-latitude trough and the light ion trough, *J. Atm. Terr. Phys.*, 44, 619.

Randall, C. E., D. E. Siskind, and R. M. Bevilacqua (2001), Stratospheric NO_x enhancements in the southern hemisphere vortex in winter/spring of 2000, *Geophys. Res. Lett.*, 28 (12), 2385–2388.

Ratcliffe, J. A. (1972), An introduction to the ionosphere and magnetosphere, *Cambridge Univ. Press.*

Rees, M. H. (1989), Physics and chemistry of the upper atmosphere, *Cambridge atmospheric and space science series, Cambridge Univ. Press.*

Remedios, J. et al. (1998), The UARS reference atmosphere project, *http://hyperion.gsfc.nasa.gov/Analysis/UARS/urap/contacts_information.html*

Richards, P. G., M. R. Torr, and D. G. Torr (1982), The seasonal effect of nitric oxide cooling on the thermospheric U.V. heat budget, *Planet. Space. Sci.*, 30(5), 515– 518.

Richards, P. G., and D. G. Torr (1984), An investigation of the consistency of the ionospheric measurements of the photoelectron flux and solar EUV flux, *J. Geophys. Res.*, 89, 5625.

Richards, P. G., J. A. Fennelly, D. G. Torr (1994), EUVAC: A solar EUV flux model for aeronomic calculations, *J. Geophys. Res.*, 99 (A5), 8981-8992.

Richardson, L. F. (1922), Weather Prediction by Numerical Process, *Cambridge Univ. Press.*

Ridley, A. J., III (1999), Variations of the thermospheric nitric oxide mass mixing ratio as a function of K_p, altitude, and magnetic local time, *Geophys. Res. Lett.*, 26 (11), 1541–1544.

Rishbeth, H., and O. K. Garriott (1969), Introduction to ionospheric physics, *Academic Press*, New York and London.

Rishbeth, H. (1972), Thermospheric winds and the F-region : A review, *J. Atmos. Terr. Phys.*, 34, 1-37.

Roble, R. G., and E. C. Ridley (1994), A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (TIME-GCM): Equinox solar cycle minimum simulations (30-500km), *Geophys. Res. Lett.*, 21, 6, 417-420.

Roble, R. (1995), Energetics of the mesosphere and thermosphere, *Upper mesosphere and lower thermosphere : A review of experiment and theory*, geophysical monograph 87, American geophysical union.

Roble, R. G., E. C. Ridley and R. E. Dickinson (1987), On the global mean structure of the thermosphere, *J. Geophys. Res.*, 92 (A8), 8745-8758.

Roble, R. G., and G. G. Shepherd (1997), An analysis of wind imaging interferometer observations of O(^{1}S) equatorial emission rates using the thermosphere-ionosphere-mesosphere electrodynamics circulation model, *J. Geophys. Res.*, 102 (A2), 2467-2474.

Rusch, D. W., and C. A. Barth (1975), Satellite measurements of nitric oxide in the polar region, *J. Geophys. Res.*, 80, 3719– 3721.

Russell, J. M., III, L. L. Gordley, J. H. Park, S. R. Drayson, D. H. Hesketh, R. J. Cicerone, A. F. Tuck, J. E. Frederick, J. E. Harries, and P. J. Crutzen (1993), The halogen occultation experiment, *J. Geophys. Res.*, 98, 10,777.

Schoendorf, J., A. D. Aylward, and R.J. Moffett (1996), Modelling high-latitude electron densities with a coupled thermosphere-ionosphere model, *Ann. Geophys.* 14, 1391–1402.

Schunk, R. W., and J. C. G. Walker (1973), Theoretical ion densities in the lower ionosphere, *Planet. Sp. Sci.* 21, 1875-1896.

Seppälä, A., P. T. Verronen, E. Kyrölä, S. Hassinen, L. Backman, A. Hauchecorne, J. L. Bertaux, and D. Fussen (2004), Solar proton events of October–November 2003: Ozone depletion in the Northern Hemisphere polar winter as seen by GOMOS/Envisat, *Geophys. Res. Lett.*, 31, L19107.

Sharma, R. D., and R. G. Roble (2002), Cooling mechanisms of the planetary thermospheres: The key role of O atom vibrational excitation of CO_2 and NO, *Chemphyschem*, 3, 841-843.

Shepherd, G. G., C. McLandress, and B. H. Solheim (1995), Tidal influence on O(^{1}S) airglow emission rate distributions at the geographic equator as observed by WINDII, *Geophys. Res. Lett.*, 22, 275.

Shepherd, G. G. (2000), The middle atmosphere, *J. Atmos. Sol. Terr. Phys*, 62, 1587-1601.

Shved, G. M., L. E. Khvorostovoskya, I. Yu Potekhin, A. I. Demiyanikov, A. A. Kutepov, and V. I. Fomichev (1991), Measurements of the quenching rate constant for collisions $CO_2(01^10)$-O : The importance of rate constant magnitude for the thermal regime and radiation of the lower thermosphere, *Atmos. Ocean. Phys.*, 27, 295-299.

Shved, G. M., A. A. Kutepov, and V. B. Ogibalov (1998), Non-local thermodynamic equilibrium in CO_2 in the middle atmosphere, I Input data and populations of the ν_3 mode manifold states, *J. Atmos. Sol. Terr. Phys.*, 60, 289-314.

Siskind, D. E., C. A. Barth, and R. G. Roble (1989a), The response of thermospheric nitric oxide to an auroral storm: 1. Low and middle latitudes, *J. Geophys. Res.*, 94 (A12), 16,885-16,898.

Siskind, D. E., C. A. Barth, D. S. Evans, and R. G. Roble (1989b), The response of thermospheric nitric oxide to an auroral storm: 2. Auroral latitudes, *J. Geophys. Res.*, 94 (A12), 16,899-16,911.

Siskind, D. E., C. A. Barth, and D. D. Cleary (1990), The possible effect of solar soft X rays on thermospheric nitric oxide, *J. Geophys. Res.*, 95(A4), 4311–4317.

Siskind, D. E., and D. W. Rusch (1992), Nitric oxide in the middle to upper thermosphere, *J. Geophys. Res.*, 97 (A3), 3209-3217.

Siskind, D. E., D. J. Strickland, R. R. Meier, T. Majeed, and F. G. Eparvier (1995), On the relationship between the solar soft X ray flux and thermospheric nitric oxide: An update with an improved photoelectron model, *J. Geophys. Res.*, 100 (A10), 19,687–19,694.

Siskind, D. E., Bacmeister, M. E. Summers, S. Zasadil, and .M. Russell III (1997), Two dimensional model calculations of middle atmosphere nitric oxide and comparisons with HALOE data, *J. Geophys. Res.*, 102 (D3), 3527-3545.

Siskind, D. E. (2000), On the coupling between middle and upper atmospheric odd nitrogen, in *Atmospheric Science Across the Stratopause, Geophys. Monogr. Ser.*, vol. 123, edited by D. E. Siskind, S. D. Eckermann, and M. E. Summers, pp. 101–116, AGU, Washington, D. C.

Siskind, D. E., J. M. Picone, M. H. Stevens, and K. Minschwaner (2004), Middle and upper thermospheric odd nitrogen: 1. A new analysis of rocket data, *J. Geophys. Res.*, 109, A01303.

Smith, F. L., and C. Smith (1972), Numerical evaluation of Chapman's grazing incidence integral ch(X,χ), *J. Geophys. Res.*, **77**, 3592-3597.

Solomon, S. C., and R. R. Garcia (1984), Transport of thermospheric NO to the upper stratosphere?, *Planet. Space. Sci.*, 32(4), 399–409.

Solomon, S., R. R. Garcia, J. J. Olivero, R. M. Bevilacqua, P. R. Schartz, R. T Clancy, and D. O. Muhleman (1985), Photochemistry and transport of carbon monoxide in the middle atmosphere, *J. Atmos. Sci.*, 42, 1072-1083.

Solomon, S. C., P. B. Hayes, and V. Abreu (1988), The auroral 6300Å emission: Observations and modeling . *J. Geophys. Res.*, 93, 9867-9882.

Solomon, S. C., and V. J. Abreu (1989), The 630-nm dayglow, *J. Geophys. Res.*, 94, 6817-6824.

Solomon, S.C., C. A. Barth, P. Axelrad, S. M. Bailey, R. Brown, R. L. Davis, et al. (1996), The Student Nitric Oxide Explorer, *Proc. SPIE, 2810*, 121.

Solomon, S. C., C. A. Barth, and S. M. Bailey (1999), Auroral production of nitric oxide measured by the SNOE satellite, *Geophys. Res. Lett.*, 26, 1259-1262.

Solomon, S. C., Auroral particle transport using Monte Carlo and hybrid methods (2001), *J. Geophys. Res*, 106 (A1), 107-116.

Solomon, S. C., S. M. Bailey, and T. N. Woods (2001), Effect of solar soft X-rays on the lower ionosphere, *Geophys. Res. Lett.*, 28, 11, 2149-2152.

Smolarkiewicz, P. K. (1983), A simple positive definite advection scheme with small implicit diffusion, *Mon. Wea. Rev.*, 11, 479-486.

Stamper R., and M. Lockwood (1999), A doubling of the Sun's coronal field during the past 100 years, *Nature* , 399, 437-439.

Strickland, D. J., and R. R. Meir (1982), A photoelectron model for the rapid computation of photoelectron ionization rates, *Mem. Rep. 5004*, Naval Res. Lab. Washington D.C.

Strobel, D. F. (1978), Parameterisation of the atmospheric heating rate from 15 to 120km due to O_2 and O_3 absorption of solar radiation, *J. Geophys. Res.*, 83, 6225.

Swaminathan, P. K., D. F Strobel, D. G. Kupp, L. Acton, R. Demajistre, J-H Yee, L. Paxton, C. Krishna Kumar, D. J. Strickland, and J. W. Duff (1998), Nitric oxide abundance in the upper mesosphere and lower thermosphere region : Roles of solar soft X-rays, suprathermal N(^{4}S) atoms, and vertical transport, *J. Geophys. Res.*, 103, (A6), 11579-11594.

Swaminathan, P. K., D. F. Strobel, L. Acton, and L. J. Paxton (2001), Model update for mesospheric/thermosphere nitric oxide, *Phys. Chem. Earth*, Part C, V. 26, 7, 533-537.

Swider, W., and M. E. Gardner (1967), On the accuracy of certain approximations for the Chapman function, *Environmental Research Papers* No.272, Air Force Cambridge Research, Bedford, MA, USA.

Summers, M. E., D. E. Siskind, J. T. Bacmeister, S. E. Zasadil, D. F. Strobel, (1997), Seasonal variation of middle atmospheric distribution of CH_4 and H_2O using a new chemical-dynamical model, *J. Geophys. Res.*, 102, 3503-3526.

Tobiska, K. W., T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottoman, and O. R. White (2000), The SOLAR2000 empirical solar irradiance model and forecast tool, *J. Atmos. Sol. Terr. Phys.*, 62, 1233-1250.

Tobiska, K. W. (2001), Validating the Solar EUV Proxy, E10.7, *J. Geophys. Res.*, 106 (A12), 29,969-29,978.

Torr, M. R., D. G. Torr, R. A. Ong, and H. E. Hinteregger (1979), Ionization frequency – I. Ionization frequencies for major thermospheric constituents as a function of solar cycle 21, *Geophys. Res. Lett.*, 6, 771.

Torr, M. R., P. G. Richards, and D. G. Torr (1980a), A new determination of ultraviolet heating efficiency in the thermosphere, *J. Geophys. Res.*, 85, 6819.

Torr, M. R., P. G. Richards, and D. G. Torr (1980b), The solar ultraviolet heating efficiency in the mid-latitude thermosphere, *Geophys. Res. Lett.*, 6, 673.

Torr, M. R., D. G. Torr and H. E. Hinteregger (1980c), Solar flux variability in the Schumann-Runge continuum as a function of solar cycle 21, *J. Geophys. Res.*, 85, 6063.

Vejby-Christensen, L., D. Kella, H. B. Pedersen, and L. H. Anderson (1998), Dissociative recombination of NO+, *Phys. Rev.* A, 57, 3627.

Vitt, F. M, T. E Cravens, and C. H. Jackman (2000), A two-dimensional model of thermospheric nitric oxide sources and their contributions to the middle atmospheric chemical balance, *J. Atmos. Sol. Terr. Phys*, 62, 653-667.

Volland, H (1988) Atmospheric tidal and planetary waves, *Klumer Academic Publishers*, Dordrecht.

Ward, W. E., D. Y. Wang, B. H. Solheim, and G. C. Shepherd (1996), Observations of the two-day wave in WINDII data during January 1993, *Geophys. Res. Lett.*, 23, 2923-2926.

Ward, W. E. (1998), Tidal mechanisms of dynamical influence on oxygen recombination airglow in the mesosphere and lower thermosphere, *Adv. Space Res.*, 21, 795–806.

Ward, W. E. (1999), A simple model of diurnal variations in the mesospheric oxygen nightglow, *Geophys. Res. Lett.*, 26, 23, 3565-3568.

Wells, G. D., A. S. Rodger, R. J. Moffett, G. J. Bailey, and T. J. Fuller-Rowell, (1997), The effects of nitric oxide cooling and the photodissociation of molecular oxygen on the thermosphere/ionosphere system over the Argentine Islands, *Ann. Geophys.*, 15, 355-365.

Weinstock, J. (1982), Nonlinear theory of gravity waves: momentum deposition, Generalized Rayleigh friction, and diffusion, *J. Atmos. Sci.*, 35, 1022.

Weinstock, J. (1990), Saturated and unsaturated spectra of gravity waves and scale-dependant diffusion. *J. Atmos. Sci.*, 47, 2211-2225.

Weinstock, J. (1993) Lagrangian coordinates and their application to gravity wave spectra. In: Thrane, E.V., Bix, T.A., Fritts, D.C. (Eds.), *Coupling processes in the Lower and Middle atmosphere.* Kulwer Academic Publishers, London, pp. 241-260.

Wennberg, P. O., J. G. Anderson, and D. K. Weisenstein (1994), Kinetics of reactions of ground state nitrogen atoms ($^4S_{3/2}$) with NO and NO_2, *J. Geophys. Res.*, 99 (D9), 18,839–18,846.

Wilke, C. R. (1950), *Chemical engineering Progress*, 46, 95-104.

Woods, T. N., F. G. Eparvier, J. Fontenla, J. Harder, G. Kopp, W. E. McClintock, G. Rottman, B. Smiley, and M. Snow (2004), Solar irradiance variability during the October 2003 solar storm period, *Geophys. Res. Lett.*, 31, L10802.

Yee, J-H, G. Crowley, R. G. Roble, W. R. Skinner, M. D. Burrage, and P. B. Hays (1997), Global simulations and observations of $O(^1S)$, $O_2(^1\Sigma)$ and OH mesospheric nightglow emissions, *J. Geophys. Res.*, 102 (A9), 19949-19968.

Yudin, V. A., B. V. Khattatov, M. A. Geller, D. A. Ortland, C. McLandress, and G. G. Shepherd (1997), Thermal tides and studies to tune the mechanistic tidal model using UARS observations, *Ann. Geophys.*, 15, 1205-1220.

Zhang, S. P., R. G. Roble, and G. G. Shepherd (2001), Tidal influence on the oxygen and hydroxyl nightglows: Wind Imaging Interferometer observations and thermosphere/ionosphere/mesosphere electrodynamics general circulation model, *J. Geophys. Res.*, 106 (A10), 21,381-21,394.

Zhou, Q. H., M. P. Sulzer, and C. A. Tepley, (1997), An analysis of tidal and planetary waves in the neutral winds and temperatures observed at the E-region, *J. Geophys. Res.*, 102 (A6), 11491-11505.

Zhu, X., P. K. Swaminathan, J. H. Yee, D. F. Strobel, and D. Anderson (1997a), A globally balanced two-dimensional middle atmosphere model. Dynamical studies of mesopause meridional circulation and stratosphere-mesosphere exchange, *J. Geophys. Res.*, 102 (D11), 13,095-13,112.

Zhu, X., J. Yee, S. A. Lloyd, and D. F. Strobel (1999), Numerical modeling of chemical-dynamical coupling in the upper stratosphere and mesosphere, *J. Geophys. Res.*, 104 (D19), 23,995–24,012.

ACKNOWLEDGEMENTS

Many people have contributed to and supported the work presented in this thesis. I wish to thank all the members of APL who have made my time here an enjoyable one, particularly Dr Tracy Moffat for introducing me the art of stress relief through martial arts, Elaina Ford for keeping me such good company in Finland, and Dr George Millward for his endless enthusiasm and lust for life. I'd also like to thank Dr Matt Harris for all the help and advice he has given me, and for his continued interest in my work. Special thanks must go to Prof Alan Aylward for being my supervisor and for giving me the chance to come back into academia after several years in industry, a choice I have never regretted.

The support of my friends and family has helped me through many moments of doubt over the last few years. I would especially like to thank Sarah, Helen, Lucy and Jay for always brightening my life and for helping me see things with a fresh perspective. Gemma's sixth sense for knowing exactly when I need to be dragged from my desk will always be appreciated. Thanks also to my parents for giving me every opportunity in life and encouraging me in whatever I do.

Most of all I wish to thank Iain for his love, support and undying confidence in me, without whom I probably would have never begun, and certainly never would have finished this PhD.